



Criterios científicos, logísticos y operativos para la
realización de los Conteos Rápidos y protocolo
para la selección de las muestras
COTECORA 2020-2021

CONTENIDO

1.	Antecedentes	4
2.	Criterios científicos	6
2.1.	Elección de Diputaciones Federales	6
2.1.1.	Estratificación y tamaño de muestra	7
2.1.2.	Procedimientos de estimación	8
2.1.3.	Integración de las estimaciones	9
2.2.	Elecciones de Gubernatura	10
2.2.1.	Estratificación y tamaño de muestra	11
2.2.2.	Procedimientos de estimación	13
2.2.3.	Integración de las estimaciones	13
3.	Selección y resguardo de las muestras	15
3.1.	Consideraciones generales para la selección de las muestras.....	15
3.2.	Protocolo de selección y resguardo de las muestras	15
3.2.1.	Instalación.....	15
3.2.2.	Selección de las muestras	16
3.2.3.	Resguardo de las muestras	16
4.	Criterios logísticos y operativos	18
4.1.	Objetivos de los criterios logísticos y operativos	18
4.2.	Esquema general de funcionamiento del Conteo Rápido	19
4.2.1.	Etapas de planeación	22

4.2.2. Preparación	28
4.2.3. Ejecución del operativo de campo en la Jornada Electoral	33
4.2.4. Esquema de contingencia.....	37
4.2.5. Esquema de seguimiento	39
4.2.6. Actividades posteriores	41
5. Anexos. Diseño Muestral y Métodos de Estimación	42
5.1. Elección de Diputaciones Federales.....	42
5.1.1. Estratificación y tamaño de muestra	42
5.1.2. Procedimiento de estimación.....	60
5.2. Elecciones de Gobernatura.....	67
5.2.1. Estratificación y tamaño de muestra	67
5.2.2. Procedimiento de estimación.....	156

1. Antecedentes

El 4 de abril de 2014, rindieron propuesta constitucional el Consejero Presidente, las y los Consejeros Electorales, con lo que se integró el Consejo General, dando inicio los trabajos del Instituto Nacional Electoral (INE).

El 23 de mayo de 2014, se publicó en el DOF, el Decreto por el que se expide la Ley General de Instituciones y Procedimientos Electorales (LGIFE) y se reformaron y adicionaron diversas disposiciones en la materia, mismo que abrogó el Código Federal de Instituciones y Procedimientos Electorales.

El 7 de septiembre de 2016, el Consejo General, mediante Acuerdo **INE/CG661/2016**, emitió el Reglamento de Elecciones del Instituto Nacional Electoral (RE), al que se adicionaron las modificaciones en cumplimiento a la sentencia de la Sala Superior del Tribunal Electoral del Poder Judicial de la Federación (TEPJF), SUP-RAP-460/2016 y acumulados, y lo aprobado mediante acuerdos del Consejo General **INE/CG391/2017**, **INE/CG565/2017**, **INE/CG111/2018**, **INE/CG32/2019** e **INE/CG164/2020**.

El 8 de julio de 2020, el Consejo General, mediante Acuerdo **INE/CG164/2020**, aprobó modificaciones a diversas disposiciones del RE; en particular, los artículos 356, 357, 362, 375, 379, 381 y 384, relativos al ámbito operativo del Conteo Rápido.

El 7 de agosto de 2020, el Consejo General, mediante Acuerdo **INE/CG188/2020**, aprobó el Plan Integral y los Calendarios de Coordinación de los Procesos Electorales Locales Concurrentes con el Federal 2020-2021, en el que se contemplan, entre otras, las actividades relacionadas con los Conteos Rápidos para las elecciones de Gubernatura en las entidades con PEL 2020-2021.

El 26 de agosto de 2020, mediante Acuerdo **INE/CG218/2020**, el Consejo General aprobó el Plan Integral y Calendario del Proceso Electoral Federal (PlyCPEF) 2020-2021, a propuesta de la Junta General Ejecutiva del Instituto, en el que se contemplan, entre otras, las actividades relacionadas con el Conteo Rápido para el PEF 2020-2021.

El 7 de septiembre de 2020, en sesión extraordinaria del Consejo General, se dio inicio al PEF 2020-2021. En ese sentido, a partir del 1º de septiembre de 2020 y hasta el 16 de enero de 2021, darán comienzo los PEL Concurrentes con el Federal 2020-2021, en las respectivas Entidades Federativas.

El 6 de noviembre de 2020, mediante Acuerdo **INE/CG558/2020**, el Consejo General determinó la realización del Conteo Rápido para la elección ordinaria de Diputaciones federales por el principio de mayoría relativa, a fin de conocer las tendencias de los resultados de la votación del día de la Jornada Electoral del PEF 2020-2021.

El 6 de noviembre de 2020, el Consejo General aprobó, a través de la Resolución **INE/CG559/2020**, ejercer la facultad de Asunción parcial e implementar el Conteo Rápido en las elecciones de Gobernador o Gobernadora en los estados de Baja California, Baja California Sur, Campeche, Chihuahua, Colima, Guerrero, Michoacán, Nayarit, Nuevo León, Querétaro, San Luis Potosí, Sinaloa, Sonora, Tlaxcala y Zacatecas, durante sus PEL 2020-2021.

El 6 de noviembre de 2020, en sesión ordinaria, el Consejo General, mediante el Acuerdo **INE/CG560/2020**, aprobó la creación e integración del Comité Técnico Asesor de los Conteos Rápidos para los procesos PEF y PEL 2020-2021 (**COTECORA**).

2. Criterios científicos

Los criterios científicos del Conteo Rápido son todos los procedimientos que, con base en la probabilidad y en la estadística, se usarán para estimar, por un lado, el porcentaje de la votación válida emitida y el número de integrantes por partido político/candidatura independiente que conformarán la Cámara de Diputados, y por otro lado, el porcentaje de votos a favor de cada una de las candidaturas a las Gubernaturas, así como para estimar el porcentaje de ciudadanos que acudan a votar¹.

Debido a que se realizarán dieciséis Conteos Rápidos, uno federal y quince locales, las Asesoras y Asesores Técnicos distribuyeron las actividades, de manera que conformaron cuatro equipos de trabajo, tres equipos son responsables de cinco de los estados, en el sentido de establecer la definición del diseño muestral y realizar las estimaciones. En el caso del Conteo Rápido Federal, un equipo de dos integrantes del comité realizaría el diseño muestral que emplearán para realizar las estimaciones de las tendencias de la conformación de la cámara baja del Poder Legislativo.

2.1. Elección de Diputaciones Federales

Derivado de los acuerdos adoptados en la segunda reunión de trabajo, el COTECORA definió a los asesores que realizarán la estimación de Diputaciones Federales, designando como responsables al Dr. Carlos Erwin Rodríguez Hernández-Vela y al Dr. Luis Enrique Nieto Barajas.

Cada uno de los asesores realizará una estimación de la elección Federal, estas dos estimaciones estarán basadas en la misma información muestral y deberán realizarse bajo las siguientes consideraciones:

¹ De acuerdo con el Artículo 356, numeral 1, del RE del INE, el Conteo Rápido se define como "el procedimiento estadístico diseñado con la finalidad de estimar con oportunidad las tendencias de los resultados finales de una elección, a partir de una muestra probabilística de resultados de actas de escrutinio y cómputo de las casillas electorales, o en su caso de los Cuadernillos para hacer las operaciones de escrutinio y cómputo de casilla, cuyo tamaño y composición se establecen previamente, de acuerdo a un esquema de selección específico de una elección determinada, y cuyas conclusiones se presentan la noche de la Jornada Electoral."

- Las estimaciones se basarán en los datos de las hojas del cuadernillo para hacer las operaciones del escrutinio y cómputo de las casillas en muestra.
- Las muestras se diseñarán con al menos una confianza del 95 por ciento y una precisión tal que genere certidumbre estadística para cumplir con el objetivo.
- En las entidades federativas con elección de Gubernatura, la muestra que se utilizará para la elección de Diputaciones Federales será un subconjunto de la muestra de la entidad.
- Los resultados de dichas estimaciones se darán mediante intervalos.
- El reporte de resultados que se presentará y será difundido la noche de la Jornada Electoral especificará las condiciones bajo las cuales se obtuvieron las estimaciones y las conclusiones que de ellas puedan derivarse.

Considerando el Artículo 373 del RE donde se establece que la muestra debe abarcar la mayor dispersión geográfica posible, se determina que el diseño muestral sea estratificado, donde al interior de cada estrato se seleccionarán casillas mediante un muestreo aleatorio simple sin reemplazo.

2.1.1. Estratificación y tamaño de muestra

Para decidir cuál es la mejor estratificación se deben considerar factores como la precisión (error de estimación) que se pretende obtener, el tamaño de muestra y el nivel de confianza deseado en las estimaciones.

La definición de los estratos para el Conteo Rápido Federal se basa en el análisis de los resultados de las elecciones federales de 2012, 2015 y 2018. A partir de diferentes escenarios de tamaño de muestra y considerando Distritos Federales como criterio de estratificación, se determinó que éstos constituyen la estratificación más conveniente para alcanzar un margen de error de 0.05 puntos porcentuales al estimar la Votación Válida Emitida por partido y entre 6 y 7 escaños mal asignados para el partido con el mayor margen de error al estimar la conformación de la Cámara de Diputados, con un nivel de confianza de 95% y un tamaño de muestra de 6,000 casillas (véase anexo).

Se requiere hacer estimaciones a dos niveles: primero, a nivel de cada uno de los 300 distritos federales, para los diputados elegidos por mayoría relativa; y en segundo lugar, a nivel nacional para la determinación de los 200 diputados por

representación proporcional. Para el primer caso, se necesita tener representación muestral a nivel distrito federal, y para el segundo, se requiere representación a nivel nacional.

Debido a los diferentes husos horarios que existen en el país y a que en el Conteo Rápido de 2015 se observó una muy baja recepción de casillas en un estado, así como la dificultad que esto implica para recibir información de manera oportuna, se tomará una sobre muestra de 10 casillas por distrito federal en los estados que tienen dos horas de retraso en relación con la hora del centro y para el estado con baja recepción; y una sobre muestra de 5 casillas por distrito federal en los estados con una hora de diferencia respecto de la hora del centro, de tal manera que el número de casillas en muestra será de 6,300.

TAMAÑO DE MUESTRA			CRITERIO DE ESTRATIFICACIÓN
TOTAL	CALCULADO	POR HUSO HORARIO Y BAJA RECEPCIÓN ^{a/}	
6,300	6,000	300	Distritos federales

a/ 70 casillas corresponden a 7 distritos de Sonora, 80 a 8 distritos de Baja California, 90 a 9 distritos de Guerrero, 35 a 7 distritos de Sinaloa, 15 a 3 distritos de Nayarit y 10 a 2 distritos de Baja California Sur.

2.1.2. Procedimientos de estimación

Los enfoques estadísticos que se usarán para la estimación de los resultados de la elección federal son el clásico y el bayesiano. Cabe mencionar que, dentro de cada uno de estos enfoques, las definiciones específicas de cada uno de los elementos que los integran dan lugar a diferentes procedimientos metodológicos.

En total se usarán dos procedimientos de estimación que tendrán el mismo objetivo, no se esperan resultados idénticos, pero sí muy similares. Las estimaciones se integrarán para presentar intervalos consolidados por la unión de ellas, por lo que se obtendrá una sola estimación de la conformación de la Cámara de Diputados que cubrirá la conformación real al menos al 95% de confianza. Se seguirá el mismo procedimiento para consolidar las estimaciones para los porcentajes de la Votación Válida Emitida y de la participación.

De manera general, los procedimientos de estimación del Conteo Rápido Federal son:

ASESORES	ENFOQUE ESTADÍSTICO	TIPO DE ESTIMADOR/MODELO DE ESTIMACIÓN	FORMA DE ESTIMACIÓN DEL INTERVALO
Dr. Carlos Erwin Rodríguez Hernández-Vela	Clásico	Estimador de razón combinado para muestreo estratificado	Re-muestreo Bootstrap
Dr. Luis Enrique Nieto Barajas	Bayesiano	Método estratificado basado en un modelo normal	Con la distribución posterior

2.1.3. Integración de las estimaciones

Cada uno de los integrantes del equipo calculará un intervalo de confianza/probabilidad para cada uno de los partidos políticos/candidaturas independientes, así como la conformación de la Cámara de Diputados.

Para emitir estimaciones únicas se construirán intervalos consolidados por la unión de ellos, con base en las estimaciones de los dos especialistas.

En el anexo de este documento se presentan detalladamente los procedimientos de estimación.

2.2. Elecciones de Gubernatura

Derivado de los acuerdos adoptados en la segunda reunión de trabajo se determinó que el COTECORA conformaría tres equipos de trabajo, organizados de la siguiente manera:

- Equipo 1: Dr. Alberto Alonso, Dr. Gabriel Núñez y Mtra. Claudia Ortiz.
- Equipo 2: Dra. Michelle Anzarut, Dr. Manuel Mendoza y Mtra. Teresa Ortiz.
- Equipo 3: Mtra. Adriana Ducoing, Mtra. Patricia Romero y Dr. Raúl Rueda.

Asimismo, se acordó que cada equipo de trabajo realizará al menos dos estimaciones para cada una de las 5 entidades, como a continuación se describe:

EQUIPO	ASESORES TÉCNICOS	CONTEO RÁPIDO QUE ATENDERÁ EL EQUIPO
1	<ul style="list-style-type: none"> • Mtra. Claudia Ortiz Guerrero • Dr. Gabriel Núñez Antonio • Dr. Alberto Alonso y Coria 	<ul style="list-style-type: none"> • Baja California Sur • Guerrero • Nuevo León • Sinaloa • Tlaxcala
2	<ul style="list-style-type: none"> • Dra. Michelle Anzarut Chacalo • Mtra. María Teresa Ortiz Mancera • Dr. Manuel Mendoza Ramírez 	<ul style="list-style-type: none"> • Colima • Chihuahua • Michoacán • Nayarit • Zacatecas
3	<ul style="list-style-type: none"> • Mtra. Adriana Ducoing Watty • Mtra. Patricia Romero Mares • Dr. Raúl Rueda Díaz del Campo 	<ul style="list-style-type: none"> • Baja California • Campeche • Querétaro • San Luis Potosí • Sonora

Las estimaciones para cada entidad estarán basadas en la misma información muestral y deberán realizarse bajo las siguientes consideraciones:

- Los resultados de dichas estimaciones se darán mediante intervalos de confianza/probabilidad.

- Las estimaciones se basarán en los datos de las hojas del cuadernillo para hacer las operaciones del escrutinio y cómputo obtenidos de las muestras de casillas.
- Las muestras se diseñarán para producir estimaciones con al menos una confianza del 95 por ciento y una precisión tal que genere certidumbre estadística para cumplir con el objetivo.
- El reporte de resultados que se presentará y será difundido la noche de la Jornada Electoral especificará las condiciones bajo las cuales se obtuvieron las estimaciones y las conclusiones que de ellas puedan derivarse.
- El porcentaje esperado de CAE con una casilla en muestra será de alrededor de 80%.

Considerando el Artículo 373 del RE, el cual determina que la muestra debe abarcar la mayor dispersión geográfica posible, se establece que el diseño muestral sea estratificado, donde al interior de cada estrato se seleccionarán casillas mediante un muestreo aleatorio simple sin reemplazo. Cabe mencionar que la estratificación y tamaño de muestra para hacer las estimaciones son las que determine el equipo responsable de la entidad.

2.2.1. Estratificación y tamaño de muestra

Con base en los ejercicios de simulación realizados para cada entidad (véase Anexo) se determinó el criterio de estratificación y tamaño de muestra a emplearse en el Conteo Rápido de cada entidad con elección de Gobernatura, como se muestra en la siguiente tabla.

ESTADO	TAMAÑO DE MUESTRA	CRITERIO DE ESTRATIFICACIÓN	% DE CAE CON UNA CASILLA	PRECISIÓN TEÓRICA ^{a/}
Baja California	667	8 Distritos Federales	76.16%	0.8%
Baja California Sur	190	Distritos Locales	70.4%	De 1.1% a 1.24%
Campeche	365	7 estratos por grupos de municipios dentro de Distritos Federales	53.65%	0.9%

ESTADO	TAMAÑO DE MUESTRA	CRITERIO DE ESTRATIFICACIÓN	% DE CAE CON UNA CASILLA	PRECISIÓN TEÓRICA ^{a/}
Chihuahua	317	Distritos Federales cruce con 2 niveles de Lista Nominal	90%	1.4% ^{b/}
Colima	190	Distritos Federales con dos niveles de índice de marginación y dos niveles de índice de votos	68%	1.3% ^{b/}
Guerrero	850	Colapso de Distritos Locales con intersección de Distritos Federales	78%	Alrededor de 1.1%
Michoacán	400	Distritos Federales con dos niveles de Lista Nominal	Al menos 88%	Menor a 1%
Nayarit	334	Distritos Federales con dos niveles de índice de marginación y dos niveles de índice de votos	64%	1.3% ^{b/}
Nuevo León	850	Colapso de Distritos Locales con intersección de Distritos Federales	79%	Alrededor de 1%
Querétaro	385	8 tipos de sección en Distritos Federales	75.26%	0.8%
San Luis Potosí	559	7 Distritos Federales	75.43%	0.8%
Sinaloa	750	Colapso de Distritos Locales con intersección de Distritos Federales	76%	Alrededor de 1%
Sonora	410	7 Distritos Federales	78.69%	0.8%
Tlaxcala	270	Intersección de Distritos Locales con Federales y algunos estratos colapsados	73.42%	De 0.87% a 1.11%
Zacatecas	376	Distritos Federales con dos niveles de índice de marginación y dos niveles de índice de votos	78%	1.3% ^{b/}

a/ Precisión esperada bajo la muestra completa con el diseño especificado para cada entidad (Ver anexos).

b/ Cuantil 0.95 de precisión sobre los 4 candidatos presidenciales de 2018, utilizando la tabla de cómputos distritales.

2.2.2. Procedimientos de estimación

En la siguiente tabla se muestra el enfoque estadístico y método de estimación con el que cada uno de los equipos determinará el intervalo de acuerdo con la sección 2.2.

ASESOR	ENFOQUE ESTADÍSTICO	TIPO DE ESTIMADOR/MODELO DE ESTIMACIÓN	MÉTODO DE CÁLCULO DE INTERVALO DE CONFIANZA
Dr. Alberto Alonso y Coria	Clásico	Estimador de razón combinado para muestreo estratificado	Métodos asintóticos
Dra. Michelle Anzarut Chacalo y Mtra. María Teresa Ortiz Mancera ^{a/}	Bayesiano	Modelo binomial negativo jerárquico	Distribución predictiva de la proporción de votos
Mtra. Adriana Ducoing Watty	Clásico	Estimador de razón combinado	Re-muestreo Bootstrap
Dr. Manuel Mendoza Ramírez	Bayesiano	Método estratificado basado en un modelo normal y ponderado con el listado nominal	Con la distribución posterior
Dr. Gabriel Núñez Antonio	Clásico	Estimador de razón combinado para muestreo estratificado	Re-muestreo Bootstrap
Mtra. Claudia Esther Ortiz Guerrero	Clásico	Estimador de razón combinado para muestreo estratificado	Re-muestreo Bootstrap
Mtra. Patria Isabel Romero Mares	Clásico	Estimador de razón combinado	Métodos asintóticos
Dr. Raúl Rueda Díaz del Campo	Bayesiano	Modelo normal	Distribución predictiva del total de votos

a/ Además de utilizar el Modelo Binomial Jerárquico realizarán estimaciones con un Estimador de Razón Combinado, que utilizarán para monitorear sus estimaciones.

2.2.3. Integración de las estimaciones

Cada uno de los miembros del COTECORA calculará un intervalo de confianza/probabilidad para cada uno de los candidatos/as contendientes de

las entidades correspondientes e integrarán sus estimaciones como se describe a continuación:

- El equipo 1 realizará tres estimaciones por entidad federativa y para cada candidato/a considerarán la unión de al menos dos de ellas. Sin embargo, la manera definitiva de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros. Los estados de los cuales son responsables son: Baja California Sur, Guerrero, Nuevo León, Sinaloa y Tlaxcala.
- El equipo 2 realizará dos estimaciones por entidad federativa, para cada candidato/a. Construirán intervalos consolidados por la unión de ambas para los estados de Colima, Chihuahua, Michoacán, Nayarit y Zacatecas.
- El equipo 3 realizará tres estimaciones por entidad federativa y reportarán un intervalo según el monitoreo de resultados de las tres estimaciones de los estados que son responsables: Baja California, Campeche, Querétaro, San Luis Potosí y Sonora.

3. Selección y resguardo de las muestras

3.1. Consideraciones generales para la selección de las muestras

Las muestras con las que se estimará, por un lado, la conformación de la Cámara de Diputados y por el otro, la votación a favor de los candidatos/as a las Gubernaturas de cada una de las entidades con elección concurrente, se obtendrán en un acto público el 4 junio de 2021. En el acto estará presente un fedatario que será testigo del desarrollo del protocolo desde la instalación del software requerido hasta la obtención y resguardo de las muestras definitivas.

Para seleccionar las muestras se hará uso de un equipo de cómputo habilitado con software estadístico. El marco muestral será el listado de las casillas aprobadas para las elecciones del 6 de junio de 2021 y la selección se realizará de acuerdo con el diseño de muestreo establecido por el COTECORA.

3.2. Protocolo de selección y resguardo de las muestras

Para la selección y resguardo de las muestras se realizarán las siguientes actividades:

3.2.1. Instalación

1. Personal de la Dirección Ejecutiva del Registro Federal de Electores (DERFE) realizará ante Fedatario la validación del equipo de cómputo donde se instalará el programa para la obtención de las muestras.
2. El COTECORA entregará el software necesario para la selección de las muestras junto con sus códigos de integridad (programa para la selección de las muestras y la base de datos de casillas, así como sus correspondientes códigos de integridad).
3. Personal de la DERFE obtendrá los códigos de integridad de los archivos entregados por el COTECORA y los comparará con los entregados. El Fedatario validará que se trata del mismo código de integridad.

4. A la vista de los asistentes se instalará el software estadístico.
5. Se solicitará a un miembro del COTECORA iniciar la etapa de ejecución de la selección de las muestras.

3.2.2. Selección de las muestras

1. Para seleccionar las muestras se requiere de un número aleatorio denominado semilla, el cual se utilizará para generar las dieciséis muestras.
2. La semilla se construirá con tres números de seis dígitos.
3. Para construir y capturar los tres números se requiere la participación de seis personas elegidas entre los asistentes. Tres, anotarán un número en un formato diseñado para tal efecto y lo ingresarán en el programa. Para asegurar el correcto ingreso, el programa solicitará la confirmación de los números, los cuales serán nuevamente ingresados por las tres personas restantes.
4. A la vista del Fedatario, los números se ingresarán en el programa de selección de las muestras. Los demás asistentes no conocerán estos números.
5. Una vez ingresados los números para construir la semilla, un miembro del COTECORA ejecutará el programa para seleccionar las muestras. Estas últimas quedarán grabadas en el disco duro.
6. Se generará un código de integridad de las muestras. El código de integridad será impreso y se entregará al Fedatario y a los asistentes que lo soliciten.

3.2.3. Resguardo de las muestras

1. Las muestras serán grabadas en dos discos compactos no regrabables.
2. Un disco será entregado al Director Ejecutivo de la DERFE para su distribución a los VOED, con el fin de que se preparen los trabajos de acopio de las hojas del cuadernillo para hacer las operaciones de escrutinio y cómputo.

3. El otro disco con las muestras, los códigos de integridad y los formatos en los que se anotarán los números para construir la semilla serán guardados en un sobre.
4. Asimismo, el equipo de cómputo en el cual fueron generadas las muestras se resguardará en un sobre.
5. Ambos sobres serán sellados y rubricados por el Fedatario.
6. El Fedatario entregará al Secretario Ejecutivo del INE o a la persona que éste último designe los sobres a los que se refiere el numeral anterior para su resguardo.
7. El Secretario Ejecutivo y el Director Ejecutivo de la DERFE, en sus respectivos ámbitos de competencia, serán responsables del resguardo y la secrecía de las muestras, hasta que se publiquen los reportes de las estimaciones de los resultados de cada elección.

4. Criterios logísticos y operativos

4.1. Objetivos de los criterios logísticos y operativos

La operación logística para el Conteo Rápido considera la definición de los recursos necesarios para planear el operativo de campo, así como de las acciones que se implementarán para asegurar el adecuado flujo de la información de las casillas electorales de la muestra, para las elecciones Federal y locales, al COTECORA el día de la Jornada Electoral. En razón de lo anterior, se busca cumplir con los siguientes objetivos:

a) General

Proveer, de manera confiable y oportuna al COTECORA la información de los resultados de las votaciones asentados en los cuadernillos de las casillas electorales de las muestras correspondientes, con la finalidad de que realice las estimaciones estadísticas para conocer las tendencias de las votaciones de la Elección Federal y de las elecciones locales de Gubernatura, el día de la Jornada Electoral.

b) Específicos

- Determinar los requerimientos para la etapa de planeación de la operación logística del Conteo Rápido.
- Precisar las funciones que desarrollará el personal involucrado en la ejecución de la operación logística del Conteo Rápido.
- Definir los procedimientos para la recopilación, reporte y captura de los datos de la votación emitida en cada una de las casillas electorales de la muestra.
- Definir el esquema de seguimiento para asegurar la oportunidad en la transmisión de los datos de las votaciones.
- Definir un esquema de contingencia que contemple soluciones ante complicaciones en el reporte de los resultados de las votaciones emitidas en cada una de las casillas electorales de la muestra.

4.2. Esquema general de funcionamiento del Conteo Rápido

Con la finalidad de brindar información oportuna a la población, tras el cierre de las casillas electorales, sobre las tendencias de los resultados de las votaciones de la Elección de diputaciones federales y de las elecciones locales de Gobernatura, el INE llevará a cabo el día de la Jornada Electoral, los ejercicios de Conteo Rápido, con base en la información asentada en los cuadernillos para hacer las operaciones de escrutinio y cómputo para casilla.

Para lograr este propósito, se requiere planear y ejecutar una serie de actividades coordinadas que tienen su base en el siguiente procedimiento general, mismo que tendrá verificación el día de la Jornada Electoral:

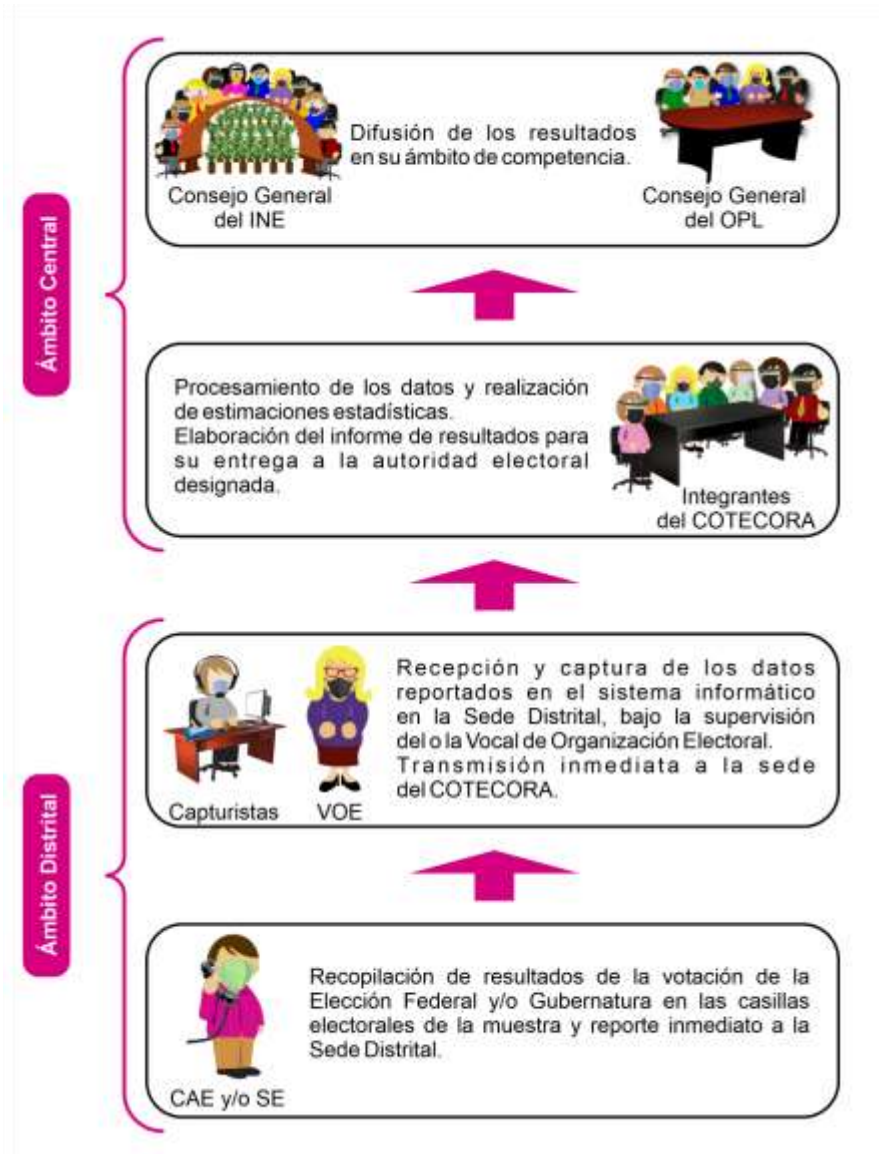
1. El personal en campo, CAE y/o SE, se encargará de recabar en el formato para la recopilación de los resultados de la votación, los resultados asentados en el cuadernillo para hacer las operaciones de escrutinio y cómputo de la(s) casilla(s) electoral(es) seleccionada(s) dentro de la muestra perteneciente(s) a su Área de Responsabilidad Electoral (ARE).
 - a. Una vez llenados los formatos, de manera inmediata llamará, a través del medio de comunicación que le fue asignado (teléfono celular, satelital o telefonía pública), a la Sede Distrital correspondiente para reportar la información recopilada.²
2. En la Sede Distrital, los y las capturistas recibirán las llamadas e ingresarán directamente en el *Sistema Informático del Conteo Rápido* los datos que le sean comunicados, esto para su transferencia inmediata a la sede del COTECORA.
3. Los integrantes del COTECORA procesarán la información proporcionada por el *Sistema Informático del Conteo Rápido* y realizarán las estimaciones estadísticas correspondientes. A partir de ello, elaborarán un informe sobre los resultados obtenidos y lo enviarán a la autoridad electoral que corresponda.

² El reporte se realizará al número telefónico del sistema multilíneas para funcionamiento de la Sala del SIJE.

4. La autoridad electoral respectiva dará a conocer a la opinión pública, la noche de la Jornada Electoral, los resultados del Conteo Rápido de la elección correspondiente a su ámbito de competencia.

El procedimiento general antes descrito, desde el acopio de datos en campo por parte de los y las CAE y/o SE, hasta la difusión de resultados en la noche de la Jornada Electoral, se representa esquemáticamente a continuación.

Figura 1 Procedimiento general del Conteo Rápido para el Proceso Electoral 2021



Al respecto, es importante hacer notar que la logística se desarrollará básicamente en el ámbito distrital con la participación fundamental de los y las Vocales de Organización Electoral (VOED), SE, CAE y capturistas.

En lo que sigue de este documento, se puntualizarán las etapas de planeación en donde se hace referencia a los diversos recursos requeridos para la ejecución

de este proyecto, así como los procedimientos que deberán realizarse antes, durante y después de la Jornada Electoral.

4.2.1. Etapas de planeación

4.2.1.1. Definición de recursos requeridos

En este apartado, se describen los recursos humanos y sus principales funciones, así como los recursos materiales y financieros necesarios en cada una de las juntas ejecutivas.

I. Recursos humanos

- Vocal de Organización Electoral Local (VOEL)
 - Con el apoyo del o la VOED, asegurará la oportuna disponibilidad y funcionamiento de todos los recursos humanos, materiales y financieros necesarios para el Conteo Rápido.
 - Supervisará la realización de la prueba de captura y los simulacros.
 - Vigilará el desarrollo del operativo de campo del Conteo Rápido durante la Jornada Electoral a efecto de garantizar el adecuado flujo en el reporte de la información.
 - Dará puntual seguimiento a los procedimientos que deban aplicarse en caso de alguna contingencia.
- Vocal Ejecutivo/a Distrital (VED)
 - Con el apoyo del o la VOED, asegurará la oportuna disponibilidad y funcionamiento de todos los recursos humanos, materiales y financieros necesarios para el Conteo Rápido.
 - Supervisará la realización de la prueba de captura y los simulacros.
 - Vigilará el desarrollo del operativo de campo del Conteo Rápido durante la Jornada Electoral a efecto de garantizar el adecuado flujo en el reporte de la información.

- Dará puntual seguimiento a los procedimientos que deban aplicarse en caso de alguna contingencia.
- Vocal de Organización Electoral Distrital (VOED)
 - Será el responsable directo de la ejecución del operativo de campo del Conteo Rápido en su ámbito de responsabilidad.
 - Brindará la capacitación correspondiente a los y las CAE, SE, y a los y las capturistas que formarán parte del proyecto.
 - Coordinará la realización de las prácticas, pruebas de captura y simulacros, así como todas aquellas actividades previstas para fortalecer el operativo de campo.
 - Dará aviso al personal en campo a quien le haya sido asignada una o más casillas electorales de la muestra para que, al término del escrutinio y cómputo de los votos el día de la Jornada Electoral, reporten los resultados de la votación a la Sede Distrital correspondiente.
 - Analizará, en su caso, si los y las SE deberán apoyar en el reporte de resultados de alguna casilla electoral de la muestra.
 - Dará seguimiento oportuno al reporte de resultados de las casillas electorales en muestra y, si es el caso, mantendrá comunicación constante con el o la SE para garantizar el flujo de la información.
 - En caso de presentarse saturación en las líneas telefónicas para el reporte de los resultados, deberá definir los mecanismos de información para indicar oportunamente al personal de campo a qué instancia comunicarse y así garantizar el correcto flujo de la información.
 - De ser necesario aplicar el plan de contingencia, será el responsable de brindar las indicaciones pertinentes al personal en campo con el objetivo de cumplir con el reporte de la información.
 - Resguardará los formatos después de la Jornada Electoral.

- Coordinador/a Distrital³
 - Auxiliará a los y las capturistas en caso de que se presenten fallas en el sistema informático.
 - Apoyará al o la VOED para dar seguimiento al reporte de las casillas electorales a través del sistema informático.
- Capturistas⁴
 - Recibirán capacitación sobre las funciones que deberán desempeñar en la ejecución del operativo de campo del Conteo Rápido.
 - Participarán en las prácticas, en la prueba de captura del sistema informático, en los simulacros y en otras actividades que favorezcan el uso del sistema informático.
 - Recibirán las llamadas del personal en campo para capturar, en el sistema informático, los datos de la votación que les sean comunicados.
 - Proporcionarán el código de confirmación del reporte de resultados a los y las CAE y/o SE, una vez que la información haya sido confirmada (de acuerdo con las precisiones de dictado indicadas en la capacitación) y validada por el sistema informático.
- Supervisores/as Electorales (SE)
 - Recibirán capacitación sobre las funciones que deberán desempeñar en la ejecución del operativo de campo del Conteo Rápido.
 - Participarán en la realización de los simulacros y en las actividades que les sean encomendadas por el o la VOED.
 - Deberán conocer cuántos y cuáles CAE bajo su responsabilidad tendrán casillas electorales pertenecientes a la muestra Federal y/o local.

³ Corresponde al o la Coordinadora Distrital de la Sala del SIJE.

⁴ Los y las capturistas que atenderán el SIJE y el Conteo Rápido.

- Verificarán que los y las CAE bajo su responsabilidad reporten la información de las casillas electorales de la muestra que les corresponda.
- Mantendrán comunicación constante con los y las CAE para conocer el avance del proceso de reporte de casillas electorales y atender, en su caso, cualquier eventualidad que impida el reporte oportuno de los datos.
- A solicitud del o la VOED, apoyarán en el acopio y reporte de datos de votación de las casillas electorales incluidas en la muestra; esto cuando algún/a CAE tenga que reportar más de una casilla electoral.
- Mantendrán comunicación constante con el o la VOED para dar seguimiento oportuno al reporte de resultados de las casillas electorales asignadas a su Zona de Responsabilidad Electoral (ZORE).
- Revisarán que los formatos recibidos se encuentren debidamente completados por el o la CAE.
- Capacitadores/as Asistentes Electorales (CAE)
 - Recibirán capacitación sobre las funciones que deberán desempeñar en la ejecución del operativo de campo del Conteo Rápido.
 - Participarán en la realización de los simulacros y en las actividades que les sean encomendadas por el o la VOED.
 - Recopilarán, en el formato correspondiente, los resultados de las votaciones de las elecciones Federal y/o de Gubernatura registrados en los cuadernillos de la(s) casilla(s) electoral de la muestra correspondiente, conforme a las indicaciones recibidas por parte del o la VOED, y los reportarán inmediatamente a la Sede Distrital.
 - Deberán asegurarse de conocer el tipo de elección que deben reportar en caso de que les haya sido asignada una casilla electoral perteneciente a la muestra de Conteo Rápido. Si el o la VOED no especificara este dato, es su responsabilidad solicitar que le precise esta información.

- Para agilizar la ejecución del escrutinio y cómputo de la casilla electoral de la muestra, podrán brindar orientación y apoyo a las y los funcionarios de la MDC en los procedimientos correspondientes⁵.
- En su caso, mantendrán comunicación permanente con su SE para reportar las causas que impidan el reporte de resultados de la votación de la(s) casilla(s) electoral(es) asignada(s), con el objetivo de buscar una solución que permita el reporte oportuno de los datos.
- Completarán, en el caso que corresponda, el cuestionario en línea que se les hará llegar a sus dispositivos móviles, para fundamentar el incumplimiento del reporte de resultados al Conteo Rápido.
- Capacitadores/as Asistentes Electorales Locales (CAEL) y Supervisores/as Electorales Locales (SEL).
 - Ante la asignación de más de una casilla a reportar por la o el CAE federal y que éstas se encuentren alejadas entre sí. En tal caso, la o el VOED priorizará principalmente el apoyo de la o el SE a cargo de la ZORE respectiva o a algún/a CAE federal cercana/o que no tenga asignada casilla electoral de la muestra, solo en casos excepcionales y previa aprobación del Consejo Distrital, las y los CAEL y SEL podrán auxiliar al CAE federal enviándole una fotografía del cuadernillo correspondiente. Se debe precisar que el responsable de transmitir la información será el CAE federal, conforme a los procedimientos definidos.
 - Una vez aprobado por el Consejo Distrital, previo análisis de la/el VOED, las y los CAEL y SEL que apoyarán a la/el CAE federal, recibirán capacitación sobre las funciones que deberán desempeñar, en su caso, en la ejecución del operativo de campo del Conteo Rápido. Esta capacitación podrá ser realizada en modalidad a distancia o presencial, y deberá incluir ejercicios sobre la toma y envío de la fotografía del cuadernillo a la o el CAE federal que corresponda. Dicha capacitación será responsabilidad de la o el VOED.

⁵ El o la CAE deberá enfatizar a la o el Presidente de la MDC la importancia de utilizar el cuadernillo para hacer las operaciones de escrutinio y cómputo para el llenado del acta correspondiente.

- Atender las demás actividades dentro de su ámbito de responsabilidad que disminuya la carga de trabajo del CAE federal para priorizar el reporte de resultados al Conteo Rápido.

II. Recursos materiales

El operativo de campo del Conteo Rápido basa su funcionamiento en la capacidad instalada para implementar el *Sistema de Información de la Jornada Electoral (SIJE)* en cada una de las juntas distritales ejecutivas.

Los recursos materiales se agrupan en dos categorías:

1. Para el reporte y recepción de la información

Son los medios de comunicación que utilizarán los y las CAE y SE desde campo, así como aquellos que emplearán los y las capturistas en cada Junta Distrital Ejecutiva:

- Los y las CAE y se utilizarán principalmente teléfonos celulares y, en menor medida, teléfonos públicos rurales y teléfonos satelitales, estos últimos principalmente en aquellas casillas electorales ubicadas en lugares alejados que no cuentan con servicio de telefonía convencional.
- Los y las capturistas dispondrán de diademas telefónicas vinculadas a un sistema multilíneas para atender las llamadas de los y las CAE y SE.
- Para la captura y transmisión de la información se instalarán equipos de cómputo con conexión a la RedINE, para el registro y transmisión inmediata de los datos a la sede del COTECORA a través del sistema informático. Se procurará que no estén muy cercanos a fin de evitar distorsiones de la información que recibirá cada capturista.

2. Para acondicionamiento

- Mobiliario. Se contará con sillas, mesas de trabajo y equipos de cómputo, preferentemente de los que se encuentran bajo el esquema de arrendamiento, en cantidad igual al número de líneas telefónicas en el sistema multilíneas.

- Plantas de energía eléctrica. Cada Junta Distrital Ejecutiva deberá disponer de este equipo para prevenir posibles eventualidades por falta de suministro de energía eléctrica.

III. Recursos financieros

Se contará con la asignación de recursos para el uso de telefonía celular y satelital, en menor medida, así como moneda fraccionaria para el pago de llamadas a través de telefonía pública rural. Estos recursos estarán destinados para la realización de simulacros y para su uso el día de la Jornada Electoral.

4.2.2. Preparación

Previo a la Jornada Electoral, se llevarán a cabo dos etapas de preparación para la ejecución del operativo de campo:

1. Capacitación de los recursos humanos.
2. Realización de prácticas, una prueba de captura y, al menos, tres simulacros.

Para ello, la DEOE se encargará de remitir oportunamente, a través del o la VOEL, los insumos y documentos de apoyo para que los haga del conocimiento de los y las VOED y coordine la realización de las actividades a que haya lugar.

4.2.2.1. Capacitación

Las actividades que se llevarán a cabo están dirigidas a la preparación y entrenamiento del personal involucrado en la ejecución del operativo de campo del Conteo Rápido. En este caso, debido a la pandemia provocada por la COVID-19 y para salvaguardar la salud, se dará prioridad a la capacitación virtual dirigida a los y las VOE de las juntas locales y distritales ejecutivas a través del Centro Virtual INE.

De manera inicial, del 8 al 26 de marzo de 2021, se llevó a cabo la capacitación a los y las VOE de las juntas locales y distritales ejecutivas, con la participación fundamental del personal de la Dirección de Planeación y Seguimiento (DPS),

encaminada a fortalecer los conocimientos que tendrán que transmitir al personal en campo, ya que esta actividad será responsabilidad del o la VOED bajo la estricta supervisión del o la VED.

Asimismo, se instrumentará un mecanismo de capacitación, específico y particular para reforzar el proceso de aprendizaje en todos los niveles de responsabilidad, con mayor énfasis en el caso de las entidades federativas donde adicionalmente se implementará el Conteo Rápido correspondiente a la Elección Local de Gubernatura.

Por otra parte, para el Segundo Taller de Capacitación dirigido a los y las SE y CAE (01 al 12 de abril de 2021) se consideró lo establecido en la Estrategia de Capacitación y Asistencia Electoral (ECAE), atendiendo a todas las medidas sanitarias establecidas. El o la VOED incluyó una explicación detallada sobre el operativo de campo a este personal, así como a capturistas y al o la Coordinadora Distrital. En dicho taller, será conveniente e importante considerar la asistencia de quienes integren el Consejo Distrital, así como del personal de las juntas locales ejecutivas, con el objetivo de que se capaciten y conozcan los procedimientos y las actividades que se implementarán para el cabal cumplimiento del proyecto, particularmente en lo referente al seguimiento del reporte de resultados.

Para ello, la DEOE, a través de la DPS, remitirá la Guía de procedimientos de la Operación Logística en el Ámbito Distrital en la que se detallarán las funciones y actividades que deberá realizar el personal involucrado en la recopilación, reporte, captura y transmisión de datos de las casillas electorales de la muestra. Asimismo, enviará material audiovisual en el que se integrarán los aspectos principales contenidos en la guía para facilitar el proceso de capacitación.

Con base en la experiencia de los procesos electorales previos en los que se han efectuado ejercicios de Conteo Rápido, **es fundamental que la totalidad de los y las CAE y SE participen en las diferentes etapas de preparación, como la capacitación y simulacros, para que cuenten con los conocimientos necesarios y suficientes sobre la operación logística que se ejecutará, con la finalidad de que estén preparados para realizar el reporte de resultados de alguna de las casillas electorales bajo su responsabilidad, en la noche de la Jornada Electoral.**

Posterior a la selección de la muestra, un día antes de la JE, previo análisis de la o el VOED, identificará si es necesario el apoyo de un/a CAEL o SEL con el envío de una fotografía del cuadernillo para hacer las operaciones de escrutinio y cómputo de casilla al CAE federal, una vez identificados las o los CAEL o SEL en

esta situación y previa aprobación del CD, la/el VOED se encargará de capacitarlas/os para que puedan apoyar correctamente en esta actividad.

En esta etapa de preparación, el o la VOED, en su carácter de coordinador/a del operativo de campo del Conteo Rápido, tendrá la responsabilidad de:

- Revisar los documentos enviados por la DEOE en los que se describen los procedimientos y actividades que se llevarán a cabo, así como los formatos que se utilizarán para la recopilación, reporte y captura de la información.
- Capacitar al o la Coordinadora Distrital, capturistas, SE, y CAE, para que cumplan a cabalidad con sus respectivas funciones.
- Capacitar al personal de la Junta Distrital Ejecutiva y quienes integran el Consejo Distrital respecto a los procedimientos del operativo de campo, con particular énfasis en el seguimiento al reporte de resultados durante los simulacros y la Jornada Electoral.
- Distribuir entre los y las CAE y SE los formatos con los que se recabará la información de la votación recibida de las casillas electorales que integrarán la muestra. Asimismo, dar a conocer los números telefónicos de la Sede Distrital y, en su caso, de las instancias de apoyo para la atención de contingencias.

4.2.2.2. Realización de prácticas, pruebas de captura y simulacros

Con el fin de recrear el contexto y situaciones que pudieran ocurrir durante la Jornada Electoral, previamente se llevarían a cabo prácticas, una prueba de captura y, al menos, tres simulacros de reporte de resultados.

a) Prácticas

Estos ejercicios tienen como objetivo general, ejercitar algunos procedimientos definidos para los y las capturistas, así como para el personal en campo (CAE y SE). Se realizaron del 22 de marzo al 09 de abril de 2021, conforme a lo siguiente:

- CAE y SE

En virtud de la cantidad de llamadas que puede atender INETEL de manera concurrente, en esta práctica **solo participaron las entidades que tienen Conteo Rápido con elección de Gubernatura.**

Se realizaron llamadas a INETEL para probar la marcación definida para el reporte de resultados a esta instancia de apoyo en caso de contingencia, debido a que durante los simulacros no todos y todas tienen la posibilidad de hacerlo.

Los y las CAE y SE realizaron una llamada a INETEL, en un periodo de ocho días, del 22 al 31 marzo de 2021, dentro del horario de las 09:00 a las 18:00 horas (hora del centro del país).

En los lineamientos que se les hizo llegar para esta actividad, se incluyó una programación de llamadas en la que se detalló el día y la hora en que los y las CAE de cada distrito electoral debían marcar. Es importante respetar la programación para evitar la saturación de las líneas en INETEL y que esta práctica se pueda desarrollar adecuadamente.

- Capturistas

Del 05 al 09 de abril de 2021, se verificaron las capacidades de los y las capturistas respecto a la velocidad y precisión en la captura de datos. Para lo anterior, es importante destacar los siguientes objetivos particulares:

- Contar con un diagnóstico objetivo sobre las capacidades de los y las capturistas respecto a la velocidad y precisión en la captura de datos (calidad en los datos).
- Establecer, por parte de los y las VOED, las medidas necesarias para mejorar la calidad del dictado de los y las CAE y el desempeño en la captura de información de los y las capturistas.

Estas prácticas se realizaron con la participación exclusiva de quienes actuarán como capturistas, bajo la coordinación y supervisión del o la VOED con apoyo del o la Coordinadora Distrital.

- b) Prueba de captura (sistema informático)

Este ejercicio persigue diferentes objetivos relacionados con el sistema informático, entre los que destacan:

- Verificar el acceso de los y las capturistas.
- Familiarizar al personal con el diseño y funcionamiento del sistema informático.
- Detectar los aspectos que pudieran ocasionar errores al ingresar los datos.
- Identificar posibles fallas para proceder a subsanarlas. Además, se busca probar las capacidades de los y las capturistas para la recepción y captura de datos a efecto de detectar posibles debilidades y corregirlas.

Esta prueba se realizó en un periodo de tres días, del 14 al 16 de abril de 2021, con la participación exclusiva de quienes actuarán como capturistas, bajo la coordinación y supervisión del o la VOED con apoyo del o la Coordinadora Distrital.

c) Simulacros

La realización de estos ejercicios se llevará a cabo en el horario de 14:00 a 18:00 horas (hora local) y tiene como objetivos:

- Implementar la ejecución de los procedimientos de reporte y transmisión de datos.
- Probar el funcionamiento de los medios de comunicación asignados a los y las CAE y SE desde campo.
- Verificar la correcta captura y transmisión de la información.
- Comprobar el funcionamiento del sistema informático.
- Verificar el esquema de seguimiento de reporte de casillas electorales al Conteo Rápido.

Todo ello, con la finalidad de detectar oportunamente cualquier posible falla en los aspectos enunciados y realizar los ajustes necesarios para garantizar el óptimo desarrollo de la operación logística el día de la Jornada Electoral.

En este sentido, se ha previsto la implementación de los siguientes simulacros:

- Primer simulacro, 29 de abril de 2021.
- Segundo simulacro, 09 de mayo de 2021.
- Tercer simulacro, 23 de mayo de 2021.

Para su ejecución, será necesario contar previamente con la relación de casillas electorales seleccionadas por el COTECORA, así como los datos de votación ficticia que deberán emplearse para que los y las CAE y SE participen en el reporte de la información conforme les instruya el o la VOED.

La DPS realizará la evaluación de la calidad de los datos capturados durante las pruebas de captura y los simulacros, con la finalidad de hacerlos del conocimiento del o la VOED para que, en su caso, tome las previsiones a que haya lugar para reforzar la capacitación de capturistas, SE y CAE, según lo considere necesario.

4.2.3. Ejecución del operativo de campo en la Jornada Electoral

Un aspecto fundamental para el operativo de campo en lo que se refiere a la identificación de las casillas en muestra, es marcar aquellas que podrían tener alguna dificultad en la recopilación de los datos. Para tal efecto se prevé realizar las siguientes acciones:

Previamente a la selección de la muestra, DEOE deberá entregar al COTECORA el listado de casillas con problemas de comunicación, a fin de que los archivos de las muestras contengan un campo en el que se identifique si la casilla tiene, o no, la problemática.

Adicionalmente, el COTECORA deberá identificar, en otro campo del archivo, a las casillas que correspondan a un CAE que tenga más de una casilla en muestra. Los archivos de las muestras se deberán generar por distrito federal para que sean enviados a los VOED.

Una vez entregada la muestra a las Vocalías de Organización Electoral de las juntas distritales ejecutivas se procederá con lo siguiente:

- El o la VOED identificará las casillas electorales de la muestra y el ARE en que se encuentren ubicadas a efecto de detectar aquellos casos en que algún/a CAE tenga en su ARE dos o más casillas electorales que reportar y

estas se encuentren alejadas entre sí. En tal caso, será necesario instruir al o la SE a cargo de la ZORE respectiva o a algún/a CAE cercano que no tenga asignada casilla electoral de la muestra, para que le brinde apoyo en estas funciones.

- En casos excepcionales se podrá solicitar el apoyo de la o el CAE o SE local, para que envíe una fotografía del cuadernillo al CAE federal para estar en condiciones de realizar el reporte correspondiente. La o el CAE solo podrá apoyarse en las y los CAEL y SEL, que hubieran sido capacitados para el Conteo Rápido.
- Durante el transcurso de la mañana del día previo a la Jornada Electoral, el o la VOED se comunicará con el personal en campo que haga uso de telefonía pública rural para hacerle saber que tiene asignada alguna casilla perteneciente a la muestra para el reporte al Conteo Rápido.
- Quienes dispongan de teléfono celular o satelital y que les corresponda reportar alguna casilla electoral de la muestra, recibirán la notificación por parte del o la VOED durante el día previo a la Jornada Electoral.

Con base en lo anterior, en el operativo de campo que se desarrollará durante la Jornada Electoral se ejecutarán tres fases principales: a) recopilación, b) reporte y c) captura-transmisión de la información. A continuación, se presentan con mayor detalle cada una de estas fases.

4.2.3.1. Recopilación de la información

Las acciones que ejecutarán los y las CAE, o en su caso SE, para el acopio de los datos de las casillas electorales de la muestra son las siguientes:

1. En la Jornada Electoral, acudirá a las 18:00 horas (hora local) a la casilla electoral seleccionada de su ARE. Ahí permanecerá para brindar el apoyo y orientación necesarios a las y los funcionarios de MDC para agilizar los procedimientos de escrutinio y cómputo de cada una de las elecciones que tengan lugar⁶.

⁶ El o la CAE deberá enfatizar a la o el Presidente de la MDC la importancia de utilizar el cuadernillo para hacer las operaciones de escrutinio y cómputo para el llenado del acta correspondiente.

2. Una vez que el o la Secretaria de la MDC concluya con el llenado de la primera columna de la sección Votos de la Elección Sacados de la Urna, del cuadernillo de la elección que corresponda reportar (diputaciones federales y Gubernatura), procederá de inmediato a transcribir los resultados en el formato de dicha elección, asegurándose de que cada uno de los datos transcritos corresponde fielmente con los del documento fuente.
3. Una vez terminada la transcripción de los datos, firmará el formato y llamará de inmediato a la Sede Distrital para realizar el reporte de los resultados. En su caso, los y las CAE que hagan uso de telefonía pública rural deberán trasladarse de inmediato al lugar donde esta se ubique para realizar el reporte.
4. Si eventualmente se le hubiera asignado otra casilla para reportar y esta estuviera próxima a la recién reportada, tendrá que trasladarse de inmediato a dicha ubicación para realizar el mismo procedimiento.
5. En aquellos casos donde a un/una CAE le corresponda realizar el reporte de resultados de la Elección Federal y de la Elección Local de Gubernatura en una misma casilla, deberá llenar el formato correspondiente a cada elección y realizar el reporte inmediato a la Sede Distrital de los resultados de la elección que primero concluya con el llenado de la primera columna del cuadernillo. Posteriormente, realizará lo propio con la elección pendiente.

En el caso que el llenado de ambos cuadernillos concluya al mismo tiempo, entonces, se reportarán los resultados de las dos elecciones en una sola llamada. Se deberá reportar en primera instancia los resultados de la Elección Federal y, enseguida, la Elección Local.

Tal como lo señala el Artículo 379, Numeral 7 del RE, “En el envío oportuno de información lo correspondiente al Conteo Rápido será siempre de primera prioridad en las tareas del CAE”.

4.2.3.2. Reporte inmediato a la Sede Distrital

Como se ha indicado, una vez que el personal en campo termine de transcribir los datos de la **primera columna de los cuadernillos** al formato, **de manera**

inmediata llamará a la Sede Distrital a que pertenezca para reportar la información recopilada.

En la Sede Distrital estará disponible un número definido de capturistas equipados/as con diademas telefónicas conectadas a un sistema multilíneas y equipos informáticos conectados a la RedINE.

Cuando se establezca comunicación, vía telefónica, con el o la capturista, los y las CAE deberán reportar los datos recabados conforme a lo siguiente:

- a) Dictará el tipo de elección (diputaciones federales o Gubernatura), los datos de identificación de la casilla y la clave de autenticación⁷ (confirmada por cuatro dígitos), con lo cual se acreditará la personalidad de quien reporta la información. Dicha clave es la llave de acceso al *Sistema Informático del conteo Rápido* para el ingreso de los datos de votación.

El dictado de los datos de identificación de la casilla y la clave de autenticación del o la CAE se realizará por cifras de dos dígitos.

- b) A la indicación del o la capturista, el o la CAE dictará los resultados de la votación que obtuvo cada partido político, coalición, candidaturas independientes (en su caso), candidato/a no registrado/a y los votos nulos, para que el o la capturista los ingrese en dicho sistema informático. Las cifras de los resultados de las votaciones de cada casilla electoral serán proporcionadas en dos ocasiones, conforme al protocolo de dictado definido.
- c) De coincidir la doble captura de datos, el sistema informático guardará la información y mostrará un código de confirmación que deberá ser dictado, por el o la capturista al o la CAE, para asegurar que cumplió con el reporte de la información. En caso contrario, el o la capturista solicitará nuevamente al o la CAE la confirmación de los datos, conforme al protocolo de dictado definido, hasta que estos coincidan.

⁷ Las claves de autenticación utilizadas durante los simulacros **serán distintas de las que se emplearán durante la Jornada Electoral**. En este sentido, para la Jornada Electoral, el o la Coordinadora Distrital deberá disponer de la relación de las claves asignadas al personal en campo en la que se incorporen datos personales para que, de ser necesario, tenga elementos para identificar la personalidad de estos en caso de que no cuenten con dicha clave.

- d) El o la CAE concluirá la llamada cuando el o la capturista se lo indique.

4.2.3.3. Captura en el sistema y transmisión al COTECORA

A continuación, se detalla el procedimiento para realizar la captura de la información en las sedes distritales.

- El o la capturista utilizará una diadema telefónica para recibir la llamada del o la CAE o SE; de esta manera podrá registrar cada dato al momento en que sea dictado.
- La captura se realizará directamente en el sistema informático en dos ocasiones. El o la CAE o SE realizará el primer dictado dígito a dígito y el segundo dictado lo efectuará en cifras completas, omitiendo los ceros a la izquierda. En caso de que hubiera discrepancias entre los registros, se deberá solicitar nuevamente el dictado de los datos respectivos, conforme al protocolo de dictado definido, para corregir la información ingresada en el sistema, hasta que los datos sean consistentes.
- Una vez que son guardados exitosamente los datos, el sistema informático generará un código de confirmación para que el o la capturista lo dicte, por cifras de dos dígitos, al o la CAE o SE, y así podrá concluir la llamada telefónica.
- El o la CAE o SE registrará en el formato la clave de confirmación, con la cual comprobará que realizó el reporte de resultados de la casilla respectiva, y deberá realizar el resguardo del formato para su entrega al o la VOED correspondiente.
- La información capturada en el sistema informático se transmitirá de forma automática a la sede del COTECORA, cuyos integrantes se encargarán de procesar la información y realizar las estimaciones estadísticas correspondientes.

4.2.4. Esquema de contingencia

Como medida preventiva a escenarios adversos que pudieran ocurrir de forma imprevista el día de la Jornada Electoral y obstaculizar el reporte de datos a las sedes distritales, se implementará el reporte de resultados de la votación a la

Junta Local Ejecutiva, como esquema de contingencia en primera instancia, y, en segunda instancia, a INETEL, para garantizar el flujo de la información y dar continuidad a las acciones que realizará el personal en campo, ante situaciones como:

- Fallas en el funcionamiento de las líneas telefónicas o suspensión del servicio.
- Fallas en el sistema informático.
- Fallas en la comunicación de la RedINE.
- Suspensión del suministro de energía eléctrica.
- Toma de instalaciones o situaciones de violencia que pongan en riesgo las actividades al interior de la Junta Distrital Ejecutiva.

Para accionar el plan de contingencia, se deberá seguir el siguiente procedimiento:

1. El o la VOED deberá realizar el reporte de la situación que se presente ante el CAU y, así mismo, notificarlo a la DPS, a través del personal previamente asignado para brindarle atención. Este procedimiento podrá realizarse vía telefónica o por correo electrónico.
2. El personal de la DPS estará en continua comunicación con el o la VOED a efecto de dar seguimiento y, en su caso, proceder a aplicar el esquema de contingencia con la participación del personal de la Junta Local Ejecutiva e INETEL.
3. En la vocalía de Organización Electoral correspondiente, se encargarán de avisar a los y las CAE y SE para que realicen el reporte de resultados a la Junta Local Ejecutiva correspondiente o a INETEL. El o la VOED se apoyará en el o la Coordinadora Distrital para hacer llegar las instrucciones a los y las CAE y SE sobre el procedimiento a seguir.
4. En caso de que no se pueda establecer la comunicación con la Junta Local Ejecutiva correspondiente después de dos intentos, deberá marcar a INETEL, y así sucesivamente hasta que la llamada se reciba en alguna de las dos instancias.
5. Si el problema se soluciona antes de que concluya el reporte de casillas electorales de la muestra, se retomarán las llamadas a la Sede Distrital

correspondiente, para lo cual el o la VOED se encargará de dar aviso al personal en campo pendiente de realizar su reporte, así como a la DPS.

6. El o la VOED tendrá la obligación de revisar el reporte de avance del *Sistema Informático del Conteo Rápido* a efecto de dar seguimiento y, en su caso, retomar la comunicación con el personal de campo para asegurar el adecuado flujo de la información.

4.2.5. Esquema de seguimiento

Con la finalidad de asegurar el adecuado y oportuno flujo en el reporte de información, se plantean diferentes mecanismos para el esquema de seguimiento a nivel distrital y central, los que se describen a continuación.

4.2.5.1. Nivel distrital

Se prevé que en el ámbito distrital se aplique un esquema de seguimiento con apoyo de los datos que proporcione el sistema informático.

Para ello, el sistema informático dispondrá de un reporte de las casillas electorales pertenecientes a la muestra, cuya información no haya sido reportada. Con base en dicho reporte, el o la VOED, con apoyo del o la Coordinadora Distrital, tendrá la obligación de verificar las casillas electorales de la elección que se encuentren en esa situación, a partir de las 20:00 horas⁸ (horario local), para que, a través de los o las SE a cargo del ARE a que estas pertenezcan⁹, se comuniquen con los y las CAE respectivos a efecto de conocer el avance en el escrutinio y cómputo o, en su caso, conocer si se presentan contratiempos que obstaculicen el reporte de resultados. De esta forma, se podrán instruir las acciones a que haya lugar a fin de garantizar el flujo de la información.

⁸ Considerando que las casillas electorales cierran en punto de las 18:00 horas y el escrutinio y cómputo de la casilla de la elección Federal dure aproximadamente 1 hora 30 minutos.

⁹ Para ello será necesario que previo a la Jornada Electoral, el o la VOED defina un esquema para estar en comunicación constante con los y las SE y estas figuras, a su vez, con los y las CAE a su cargo. Por ejemplo, si la mayoría cuenta con teléfono celular, se podría implementar un grupo de chat para una comunicación más expedita.

Una vez que se lleve a cabo la difusión de las estimaciones correspondientes a la opinión pública, los reportes de resultados que se reciban posteriormente no serán considerados por el COTECORA, motivo por el cual es fundamental que los y las VOED garanticen el mayor número de casillas electorales reportadas al Conteo Rápido antes de las 21:30 horas (hora local), o, en su defecto, hasta el momento que le sea indicado, exclusivamente, por la vía de comunicación oficial definida para el Conteo Rápido.

Una vez que la DPS notifique al o la VOED, este podrá dar aviso al personal en campo que no hubiere concretado el reporte de datos que se podrá prescindir de hacerlo; no obstante, se deberá conocer los motivos que impidieron llevar a cabo dicho reporte, conforme a las indicaciones que en su momento sean indicadas.

4.2.5.2. Nivel central

En el ámbito central se plantean los métodos de seguimiento que a continuación se enuncian:

1. La semana anterior a la Jornada Electoral, se integrará un directorio de seguimiento a nivel central en el que se incorporará el personal de la DPS, de la Dirección de Operación Regional (DOR) y los y las VOEL o, en su defecto, quien ellos designen para dar seguimiento.
2. La DPS, en apoyo con la DOR, darán seguimiento puntual al reporte de información de las casillas electorales de la muestra que se hayan reportado al COTECORA, a través del sistema informático.
3. A través de grupos de chat y/o llamadas telefónicas, en caso de ser necesario, personal de la DPS se comunicará, en primera instancia, con los y las VOEL a efecto de hacer de su conocimiento el retraso en el reporte de casillas electorales al COTECORA y conocer la problemática que acontece.
4. En casos excepcionales, cuando no se logre establecer comunicación con la Junta Local Ejecutiva, el personal de la DPS podrá comunicarse directamente con el VOED.
5. El personal de la DPS compartirá la información de las problemáticas con el titular de la DEOE del INE.

4.2.6. Actividades posteriores

Posterior a la Jornada Electoral, los y las VOED tendrán la obligación de recabar los formatos de todos y todas las CAE que les correspondió participar en el ejercicio del Conteo Rápido, con el apoyo de los y las SE. Aquellos o aquellas que sí pudieron concretar el reporte de alguna(s) casilla(s) de la muestra, deberán verificar que el formato esté debidamente llenado, firmado y con la clave de confirmación de reporte.

En el caso de los y las CAE que debían realizar el reporte de los resultados de su casilla electoral al Conteo Rápido y no lo cumplieron o lo hicieron después del horario indicado, deberán responder, entre el 8 y el 10 de junio de 2021, un cuestionario en línea a través de una liga que les llegará a su dispositivo móvil.

En este contexto, cobra relevancia el llenado de los formatos, así como su conservación, por parte de los y las CAE, posterior al reporte de resultados de la votación.

El o la VOED deberá elaborar una relación de estos documentos y conservarlos para que, una vez que sean solicitados por la DPS, los remita conforme a la forma y tiempos que le sean comunicados.

Finalmente, es importante destacar que estos documentos serán objeto de un minucioso análisis para realizar una evaluación integral sobre la operación logística de los conteos rápidos.

5. Anexos. Diseño Muestral y Métodos de Estimación

5.1. Elección de Diputaciones Federales

5.1.1. Estratificación y tamaño de muestra

5.1.1.1. Dr. Luis Enrique Nieto Barajas¹⁰

Introducción

Para el diseño muestral de la elección federal de 2021 nos basaremos en los resultados de la elección federal del 2015. Consideramos que la elección federal del 2015 presentó retos importantes que deben de ser tomados en cuenta en el diseño para el 2021. Por otro lado, la elección federal más reciente, la del 2018, fue una elección dominada por un solo partido mayoritario, por lo que presenta menos retos metodológicos.

Las elecciones federales de México de 2015, denominadas oficialmente por la autoridad electoral como el Proceso Electoral Federal 2014-2015, fueron las elecciones intermedias llevadas a cabo en México el 7 de junio de 2015 para la elección de 500 miembros de la cámara de diputados federales. Debido a la reforma electoral de 2014, en estas elecciones se eligieron simultáneamente los puestos a cargos federales y locales en 17 entidades del país.

El número de escaños obtenidos oficialmente por cada uno de los partidos políticos se presenta en la Figura 1.

Determinación del tamaño de muestra y error de estimación

Como ya se mencionó, el diseño muestral estará basado en un muestreo estratificado por distrito federal. Para la determinación del tamaño de muestra y de los errores de estimación alcanzados con cada tamaño, nos basaremos en un estudio de simulación.

¹⁰ Con la colaboración del Mtro. Carlos Samuel Pérez Pérez

Se generan 1,000 muestras aleatorias estratificadas de casillas de tamaños 3,000, 4,500, 6,000, 7,500 y 9,000, que divididas entre los 300 distritos electorales equivalen a 10, 15, 20, 25 y 30 casillas por distrito.

Para cada muestra se estiman los parámetros $\theta_{i,j}$ de manera bayesiana y se obtienen 5,000 realizaciones de la distribución posterior. Para cada una de estas realizaciones calculamos el número de escaños NE_j . Suponiendo una función de pérdida valor absoluto, estimamos NE_j mediante la mediana posterior a la que denotaremos por \widehat{NE}_j .

Definimos el error de estimación como el número de escaños mal asignados por partido $NEMA_j = |NE_j - \widehat{NE}_j|$, donde NE_j es el valor oficial de escaños obtenidos por el partido j descrito en la Figura 1. Resumimos el error de estimación en la conformación de toda la cámara mediante dos medidas: el promedio de escaños mal asignados $PEMA = \frac{1}{J-2} \sum_{j=1}^{J-2} NEMA_j$; y el máximo de escaños mal asignados $MEMA = \max_{\{j=1, \dots, J-2\}} NEMA_j$.

Los resultados del ejercicio de simulación se muestran a continuación. La Figura 2 contiene histogramas para los 1,000 valores de los errores de estimación por partido $NEMA_j$. Notamos que el partido con mayor variabilidad es el PAN con errores de hasta 7 escaños, seguido por PRD y MORENA con errores de hasta 3 y 2 escaños respectivamente. Los demás partidos presentan muy poca variabilidad con errores prácticamente de cero. Recordemos que el indicador $NEMA_j$ se basa en el estimador puntual \widehat{NE}_j ; sin embargo, en la práctica no se reportará el estimador puntual sino un intervalo de credibilidad al 95%. Comparando la distribución muestral de los errores por partido en los distintos tamaños de muestra, se aprecia un ligero aumento en la variabilidad de los errores para el tamaño de muestra de 3,000 casillas (10 casillas por distrito), en cambio para los tamaños de muestra mayores no se aprecian diferencias considerables.

Una aclaración importante aquí es el hecho de que el partido del trabajo (PT), correspondiente a $j = 5$, estaba en el límite de perder el registro, de hecho, lo perdió y luego lo recuperó. Esto hace que para algunas muestras y para algunas realizaciones de la distribución posterior de λ_5 , el PT a veces tenga o no tenga asignación de escaños por el principio de representación proporcional. En consecuencia, el indicador $NEMA_j$ para algunos partidos como el PAN, MORENA y PRD toman valores más grandes que para otros partidos.

Al promediar todos los errores de estimación de todos los partidos, la Figura 3 muestra las distribuciones muestrales (histogramas de los 1,000 valores) de los errores promedios de la cámara $PEMA$. En estas gráficas sí se observa un cambio

en la dispersión de las distribuciones al aumentar el tamaño de muestra. La distribución muestral con 3,000 casillas es mucho más plana (más dispersa) que la que se obtiene con 9,000 casillas que está muy concentrada en el valor de uno. Comparando las modas de las distribuciones, prácticamente con 4,500 casillas y hasta 9,000 la moda se encuentra en un escaño de error. En cambio, la moda cuando se tienen muestras de 3,000 casillas se encuentra compartida por los valores de uno y dos escaños de error.

Para una mejor comparación entre las distribuciones, el Cuadro 1 muestra los cuantiles del número promedio de escaños mal asignados *PEMA* para toda la cámara con distintos tamaños de muestra. Las medianas $q_{0.5}$ son de 1.54 para un tamaño de muestra de 3,000, de 1.27 para un tamaño de muestra de 4,500, de 1.18 para un tamaño de muestra de 6,000 y de 1.09 para tamaños de muestra de 7,500 y 9,000.

Por otro lado, la Figura 4 muestra las distribuciones muestrales del máximo número de escaños mal asignados *MEMA* con distintos tamaños de muestra. Al aumentar el tamaño de muestra toda la distribución se mueve hacia la izquierda, lo que indica un menor error. Para 3,000 y 4,500 casillas la moda se sitúa en 10 escaños, para los tamaños de muestra de 6,000 y 7,500 la moda es de 9 y para 9,000 casillas la moda es de 8 casillas. Los cuantiles de estas distribuciones de muestreo se incluyen en el Cuadro 2. Prácticamente para tamaños de muestra de 7,500 y 9,000 no hay diferencias.

Relación entre casillas por ARE y tamaño de muestra

En el ejercicio electoral de 2015 se reportó un total de áreas de responsabilidad electoral (ARE) de 33,042 con un total de casillas de 148,833. En este caso, cada casilla fue identificada a través de los campos Estado, Distrito Federal y Número de ARE de forma anidada. No obstante, las ARE contuvieron casillas que respondían a una división operativa diferentes a la de los distritos, por lo que para una sección podía haber más de una ARE.

Para cada una de las 1,000 muestras simuladas se calcula la distribución de casillas por ARE y se promedian. Denotamos por *PCARE* al número promedio de casillas en muestra que estarían asignadas a un ARE. Este indicador, para los distintos tamaños de muestra, se reporta en la Figura 5. Observamos que para los tamaños de muestra de 7,500 o menos, al menos 90% de los ARE con muestra tienen una sola casilla, en cambio con muestras de 9,000 el porcentaje de ARE

con una sola casilla disminuye a 88%. Adicionalmente notamos que para muestras de 7,500 o 9,000 existen ARE con hasta 4 casillas.

Conclusión

Tomando en cuenta los resultados aquí presentados, los errores de estimación en la conformación de la cámara resumidos en los indicadores *PEMA* y *MEMA*, sus distribuciones de muestreo varían muy poco con muestras de tamaños mayores o iguales a 6,000. Adicionalmente, considerando las distribuciones del número de casillas por ARE, sugerimos que el tamaño de muestra para la estimación de la conformación de la cámara para el ejercicio electoral 2020-2021 sea de 6,000 casillas totales (20 casillas por distrito federal).

Consideramos que este tamaño de muestra es muy manejable por los Capacitadores y Asistentes Electorales (CAE), por lo que se esperaba que la muestra final recolectada para hacer la estimación sea de casi el 100%.

Cuadro 1: Cuantiles de la distribución de muestreo de **PEMA**.

Tamaño de muestra	q _{0.025}	q _{0.05}	q _{0.25}	q _{0.5}	q _{0.75}	q _{0.95}	q _{0.975}
3000	0.727	0.909	1.273	1.545	1.909	2.455	2.727
4500	0.727	0.727	1.091	1.273	1.636	2.182	2.364
6000	0.545	0.636	0.977	1.182	1.455	2.000	2.273
7500	0.636	0.727	0.909	1.091	1.364	1.818	2.091
9000	0.545	0.636	0.818	1.091	1.273	1.545	1.727

Cuadro 2: Cuantiles de la distribución de muestreo de **MEMA**

Tamaño de muestra	q _{0.025}	q _{0.05}	q _{0.25}	q _{0.5}	q _{0.75}	q _{0.95}	q _{0.975}
3000	8	8	9	10	11	14	14
4500	7	7	9	9	10	12	13
6000	6	7	8	9	10	11	12
7500	6	6	8	9	9	11	11
9000	6	6	7	8	9	10	11

Figura 1 Resultados desglose

Partido	Diputados Mayoría relativa	Diputados Rep. Proporcional	Total
 Partido Revolucionario Institucional	155	48	203
 Partido Acción Nacional	56	53	109
 Partido de la Revolución Democrática	28	28	56
 Partido del Trabajo	6	0	6
 Partido Verde Ecologista de México	29	18	47
 Movimiento Ciudadano	10	15	25
 Nueva Alianza	1	10	11
 Movimiento de Regeneración Nacional	14	21	35
 Partido Encuentro Social	0	8	8
Independientes	1	0	1
Total	300	200	500

Figura 2. Distribución de muestreo de **NEMA_j**, para los 11 partidos del 2015, por tamaño de muestra

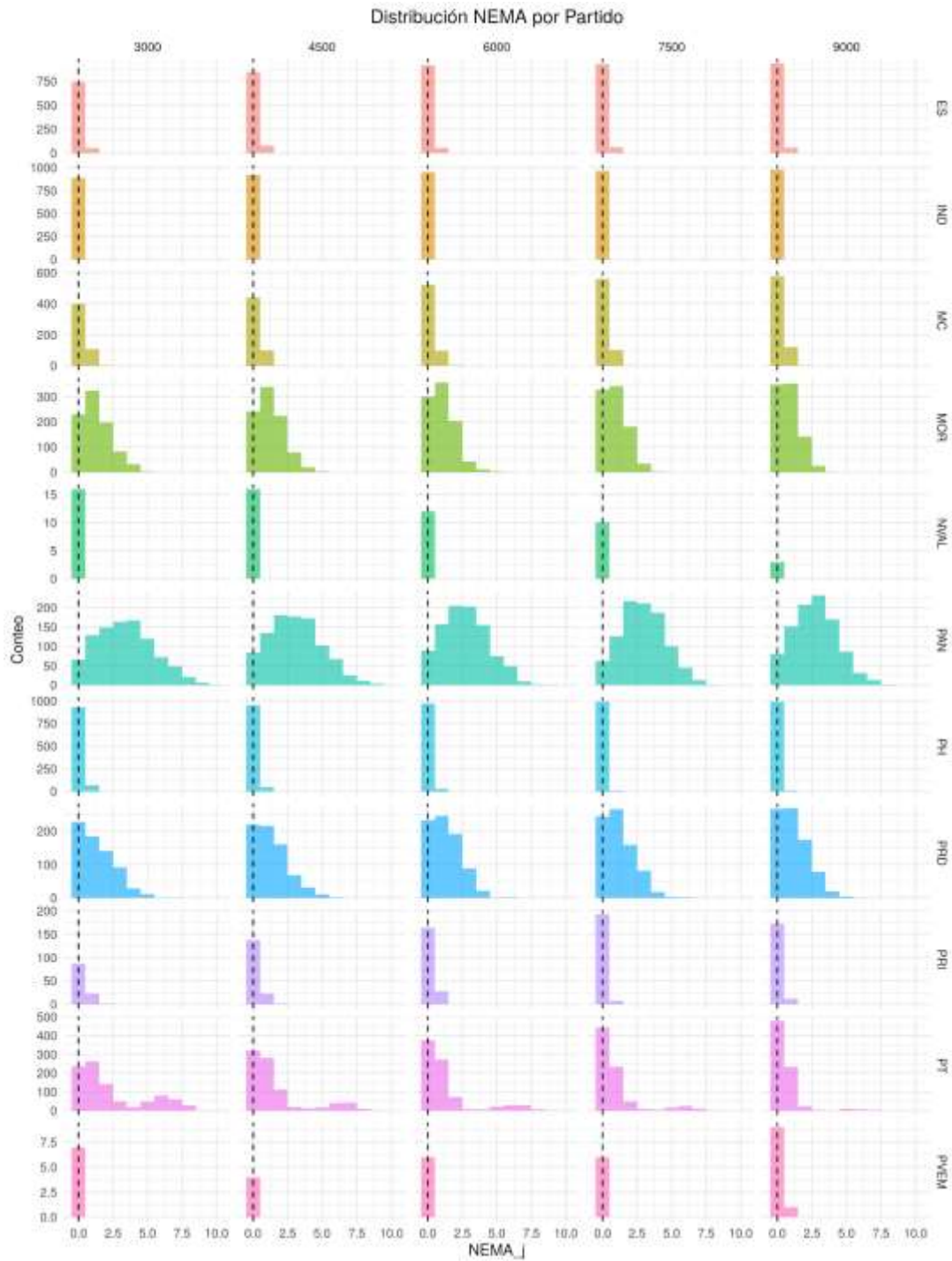


Figura 3: Distribución de muestreo de **PEMA** por tamaño de muestra.

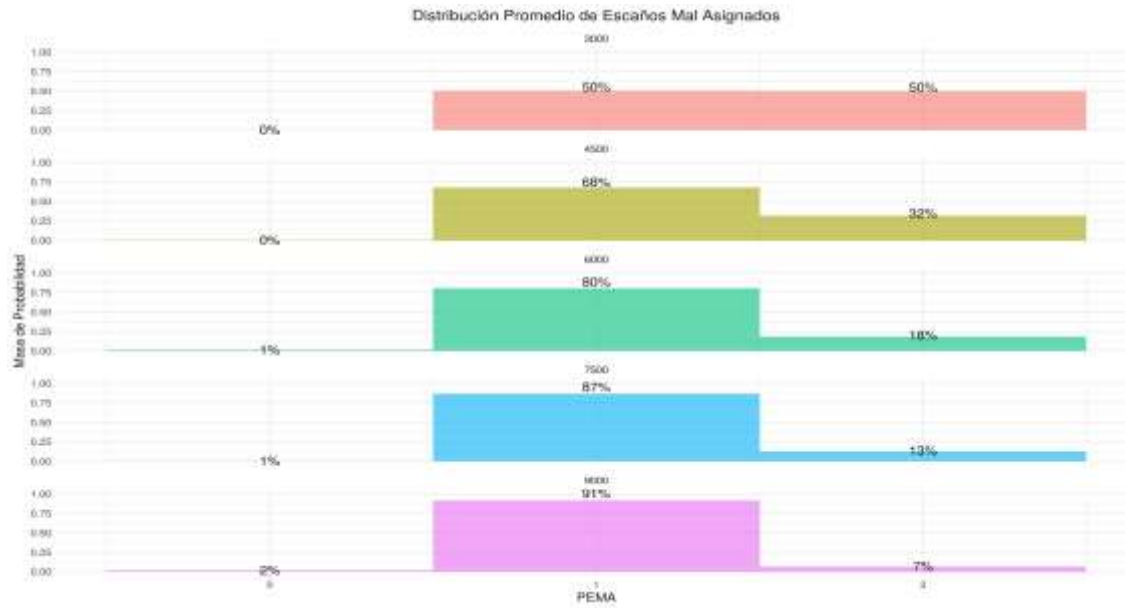


Figura 4: Distribución de muestreo de **MEMA** por tamaño de muestra.

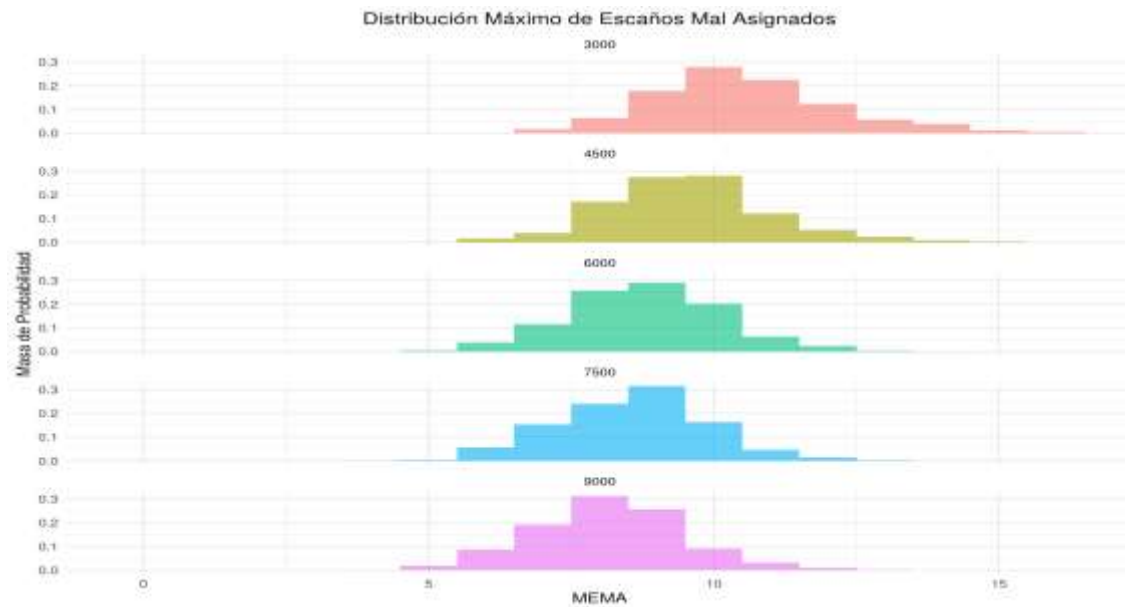


Figura 5: Distribución del número de casillas por ARE para cada tamaño de muestra.



5.1.1.2. Dr. Carlos Erwin Rodríguez Hernández-Vela¹¹

Consideraciones técnicas para el Conteo Rápido en la Elección Federal de Diputados en el proceso electoral 2021

Resumen

Para el Conteo Rápido en la Elección Federal de Diputados, se seguirá un diseño de muestreo estratificado, en donde:

- Los estratos serán los distritos federales por lo que se tendrán 300 estratos.
- El tamaño de muestra total será de 6,240 casillas a nivel nacional, en donde se seleccionarán:
 - 20 casillas en cada distrito, 6,000 casillas en total.
 - 10 casillas de sobre-muestra debido a diferencias de huso horario y a que en el Conteo Rápido de 2015 se observó una muy baja recepción de casillas en estos estados. La sobre-muestra total será de 240 casillas. Los estados con sobre-muestra son; Sonora (7 distritos – 70 casillas), Baja California (8 distritos – 80 casillas) y Guerrero (9 distritos – 90 casillas).
- En el Conteo Rápido para la Elección Federal de Diputados de 2015 se observó un nivel de respuesta del 74% y la estimación final se realizó a las 10:15 pm. Por lo tanto, se espera que de las 6,240 casillas totales para Conteo Rápido Federal de 2021 se reciba una muestra efectiva de entre 4,000 y 5,500 casillas alrededor de las 10:30 pm del 6 de junio de 2021.
- Con la muestra efectiva el margen de error en las estimaciones será de:
 - 1% al estimar la participación en la Elección Federal.
 - 0.05% al estimar la Votación Válida Emitida por partido.
 - Entre 6 y 7 escaños mal asignados para el partido con el mayor número de escaños mal asignados al estimar la conformación de la cámara de diputados.

¹¹ Con la colaboración del Act. Edgar Gerardo Alarcón González

Esta estrategia de selección y márgenes de error esperados se obtuvieron tomando como referencia los Cómputos Distritales para las elecciones de diputados en los años 2012, 2015 y 2018.

Para la estimación clásica se usarán ideas de re-muestreo Bootstrap.

Para armonizar los diseños para las elecciones locales, se consideró en el diseño que:

1. Las estratificaciones para los diseños locales son refinaciones de los distritos federales.
2. En los estados con elección local, la muestra para el conteo rápido federal será una sub-muestra de la muestra usada en el conteo rápido para la elección local.

Se realizarán dos estimaciones de la conformación de la cámara de diputados, una realizada por el equipo del Dr. Luis Enrique Nieto y el Maestro Carlos Pérez y la otra por el equipo del Dr. Carlos E. Rodríguez y el Act. Edgar Alarcón. Ambas estimaciones se consolidarán en una sola estimación mediante la unión.

Objetivos

En el Conteo Rápido para la Elección Federal del 6 de junio de 2021, se establecieron los tres objetivos siguientes:

1. Estimar la conformación de la cámara de diputados, i.e. de los 500 diputados que Federales, el número de escaños que obtiene cada partido político, así como el número de curules obtenidos por candidaturas independientes:
 - 300 diputados por mayoría relativa. El partido que alcanza la mayor votación en un determinado distrito obtiene un escaño en la cámara de diputados. Lo anterior se replica en los 300 distritos electorales en los que se divide el país. Esta vía es la única en la que un candidato/a independiente puede obtener un escaño en la cámara.
 - 200 diputados por representación proporcional. Para calcular esta parte se necesita
 - La conformación de los 300 diputados por mayoría relativa.

- La votación total emitida (VTE), la votación válida emitida (VVE) y la votación nacional emitida (VNE)¹².
 - Las reglas para distribuir escaños por representación proporcional como marcan tanto el LEGIPE como la Constitución Política de los Estados Unidos Mexicanos.
2. Estimar el porcentaje de votos obtenida por cada partido tomando como referencia la VVE. Esta votación es particularmente importante pues es la que toma como referencia la ley para determinar si un partido pierde o mantiene el registro. Si un partido iguala o supera el 3% de la VVE mantiene su registro, en otro caso lo pierde.
 3. Estimar el porcentaje de participación nacional en la elección.

Todas las estimaciones se presentarán en forma de intervalo de confianza del 95%.

Ideas básicas del muestreo probabilístico aplicadas al Conteo Rápido

Uno de los objetivos particulares del Conteo Rápido Federal es estimar el porcentaje de votos, p , con respecto a la Votación Válida Emitida que obtiene un partido con registro a nivel nacional¹³. Esta cantidad es el valor real y se obtiene una vez realizados los cómputos distritales de la elección. Sin embargo, siguiendo la teoría del muestreo probabilístico, se selecciona una muestra aleatoria de n casillas (muestra total), del total de las N casillas instaladas, y con la información recuperada n_e (muestra efectiva) en la tarde-noche del día de la elección se calcula el estimador \hat{p} . Utilizando las herramientas del muestreo, es posible definir **estrategias de selección, tamaños de muestra y estimadores** para asegurar que

$$|p - \hat{p}| \leq \epsilon, \quad (1)$$

¹² En el LEGIPE se definen tres tipos de votación: VTE = suma de todos los votos, la VVE = VTE – votos candidatos no registrados – votos nulos y VNE = VVE – votación de los partidos que no hayan rebasado el 3% de la VVE – votación candidatos independientes.

¹³ Esto se realizará para cada partido político a nivel nacional, pero para fines explicativos se describe el procedimiento sólo para un partido.

con un 95% de confianza. A la cantidad ϵ se le conoce comúnmente como **precisión, margen de error** o **error máximo aceptable** en la estimación, y se fija de acuerdo con las exigencias de la elección y con la capacidad operativa de campo. La expresión (1) se puede escribir de forma equivalente en términos de intervalos de confianza

$$\hat{p} - \epsilon \leq p \leq \hat{p} + \epsilon. \quad (2)$$

Lo anterior, se pueden interpretar tomando como referencia el siguiente ejemplo: si se extraen 100 muestras distintas e independientes una de otra (cada una siguiendo la misma estrategia de selección, usando el mismo tamaño de muestra y el mismo estimador) y con cada muestra se genera una estimación, de tal manera que se obtengan 100 estimaciones (independientes).

$$\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4, \hat{p}_5, \dots, \hat{p}_{100},$$

aproximadamente 95 estimaciones cumplirán con que su distancia a p , es menor a ϵ .

Consideraciones preliminares

De los tres objetivos establecidos para el Conteo Rápido Federal, la estimación de la conformación de la cámara de diputados es el de mayor complejidad, además

- Contiene a los otros dos objetivos de manera natural.
- Estimar la conformación de la cámara de diputados no es un problema estándar y la teoría del muestreo, tal como se vio en la sección anterior, no se puede aplicar directamente.

Lo que se busca en esta sección es introducir un criterio análogo al descrito en la expresión (1), pero para la conformación de la cámara. Primero es necesario definir una medida de error. Con esta meta se presenta la Tabla 1, que contiene lo sucedido en el Conteo Rápido de 2003.

Tabla 1. Conformación de la cámara de diputados en la Elección Federal de 2003: Estimaciones puntuales de los tres equipos

No.	Partidos	Resultado Final (12 de julio, 2003)	Estimación puntual de la conformación de la Cámara de Diputados 2003 (6 de julio, 2003)		
			Ignacio y Hortensia	Raúl y Paty	Federico y Karim
1	PAN	153	155	152	155
2	PRI	226	223	226	224
3	PRD	95	96	96	95
4	PVEM	15	15	15	15
5	PT	6	6	6	6
6	Convergencia	5	5	5	5

Siguiendo la idea mostrada en la expresión (1), para el caso de la estimación de la conformación de la cámara definiremos el error en la estimación como el número máximo de escaños mal asignados. Formalmente lo definimos como

$$\max_{j \in \{1,2,\dots,k\}} |NE_j - \widehat{NE}_j|. \quad (3)$$

En donde

- k Número de partidos registrados ante el INE a nivel nacional
- NE_j Número de escaños en la cámara de diputados asignados al partido j , una vez concluido el cómputo distrital de votos (final).
- \widehat{NE}_j Número de escaños en la cámara de diputados asignados al partido j , obtenida mediante la estimación del conteo rápido.

El número de escaños mal asignados, para el partido con el mayor número de escaños mal asignados, es una de muchas posibilidades para la medida de error. Es posible definir otros criterios de forma similar. Sin embargo, el criterio (3) es fácil de comprender y en los ejercicios realizados se observaron buenos resultados. En la Tabla 2 se muestra 1) el número de escaños mal asignados para cada fuerza política y el máximo número de escaños mal asignados de cada uno de los tres equipos en el Conteo Rápido de 2003 (Véase también la Tabla 1).

Tabla 2. Error en la estimación de la conformación de la cámara de diputados en la Elección Federal de 2003

Partidos	Escaños mal asignados Ignacio y Hortensia	Escaños mal asignados Raúl y Paty	Escaños mal asignados Federico y Karim
PAN	2	1	2
PRI	1	2	0
PRD	1	1	0
PVEM	2	2	2
PT	0	0	0
Convergencia	0	0	0
Máximo número de escaños mal asignados	2	2	2

Para definir el tamaño de muestra en la Elección Federal, primero se establecerá un máximo error permisible en la asignación de escaños, denotado como ***d***, y a continuación se buscará determinar el tamaño de muestra necesario para garantizar que

$$\max_{j \in \{1,2,\dots,k\}} |NE_j - \widehat{NE}_j| \leq d,$$

con un 95% de confianza.

No es posible encontrar expresiones analíticas que permitan despejar el tamaño de muestra necesario para alcanzar un margen de error ***d*** este caso. Sin embargo, mediante ejercicios de simulación es posible aproximar la distribución para (3) usando varios tamaños de muestra.

Ejercicios de simulación

Para definir el diseño de muestreo para el Conteo Rápido 2021 Se usaron las bases de datos de referencia con los resultados de los cómputos distritales en las elecciones de diputados 2012, 2015 y 2018.

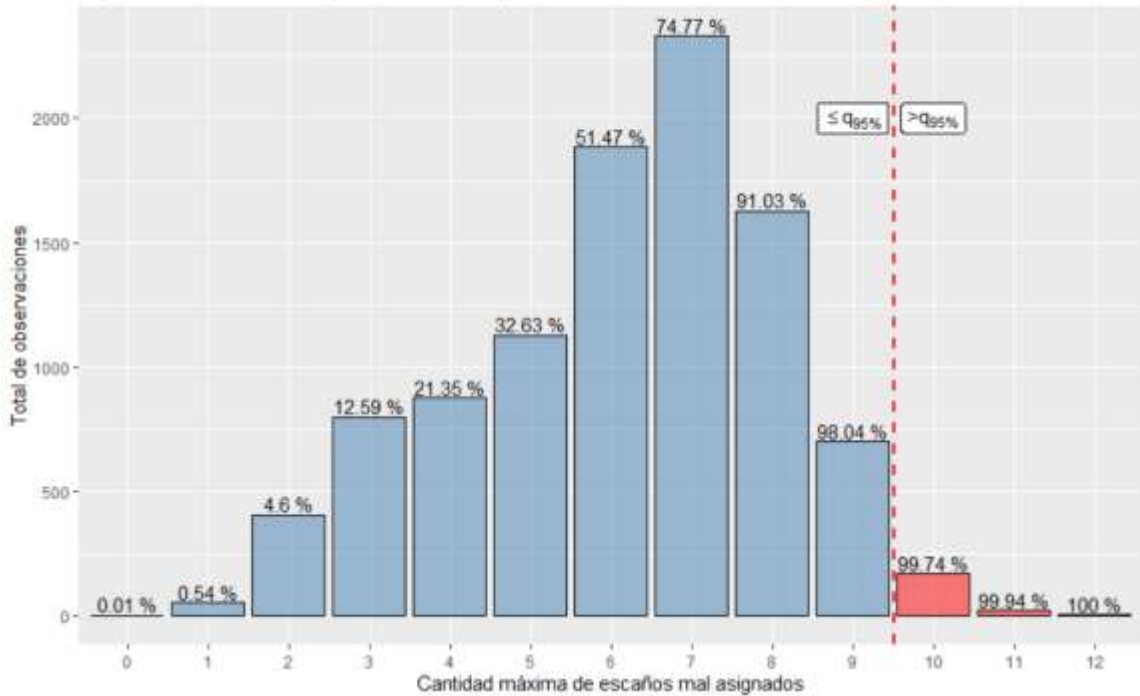
En todos los experimentos se siguió un diseño de muestreo estratificado por distrito federal electoral en donde, en cada iteración se siguieron los siguientes pasos;

- Selección por MASSR de n_h casillas al interior de cada distrito, considerando una de las bases de referencia.

- Se usó el estimador común del total para aproximar el total de votos en favor de cada fuerza política.
- Considerando las coaliciones y los votos totales en cada distrito se obtuvieron los diputados por el principio de mayoría relativa.
- Se usó la teoría del muestreo aleatorio estratificado para combinar los resultados de los 300 distritos y estimar los distintos tipos de votaciones (votación total emitida, votación válida emitida y votación nacional emitida).
- Utilizando el estimador de razón combinado se realizó la estimación de la participación, así como del porcentaje de la VVE que alcanzó cada partido político.
- Se estimó la conformación de la cámara de diputados.
 - En este punto ya se cuenta con estimaciones para los diputados por MR, así como para la VTE y VVE.
 - Se siguieron las reglas marcadas en la LEGIPE y en la Constitución para distribuir escaños por representación proporcional. Ver Apéndice.

Lo anterior constituye una realización de la simulación, se realizaron 10,000 iteraciones y en cada iteración se calculó el margen de error mostrado en la expresión (3). Se probaron los tamaños de muestra al interior de cada estrato de $n_h = 5, 10, 15, 20, 25$ y 30 , dado que son 300 estratos y en cada estrato se usará el mismo tamaño de muestra, lo anterior implica que los tamaños de muestra a considerar son de $n = 1,500, 3,000, 4,500, 6,000, 7,500$ y $9,000$ casillas. En la Figura 1 se muestra la distribución de muestreo para el número máximo de escaños mal asignados, utilizando como referencia la elección de 2015 y $n_h = 20$ ($n = 6,000$).

Figura 1: Densidad y distribución acumulada estimada del número máximo de escaños mal asignados
Experimento realizado con 10,000 muestras independientes



En la Tabla 3 se presentan los resultados de las simulaciones, considerando todas las votaciones de referencia y los distintos tamaños de muestra.

Tabla 3. Margen de error para la estimación de la conformación de la cámara de diputados vía un muestreo estratificado, considerando los cómputos distritales 2012, 2013 y 2018

n_h	2012	2015	2018
5	9	10	9
10	7	9	7
15	6	9	6
20	6	9	6
25	5	8	5
30	5	8	5

Armonización entre los diseños estatales y el diseño federal

En el Conteo Rápido 2021, se estimarán 15 gubernaturas, así como la conformación de la Cámara de Diputados. Por tal motivo, en los estados en donde habrá elección a Gobernador no se seguirá un diseño estratificado por Distrito Federal. De manera general podemos decir que en los estados en donde haya elecciones para gobernador se usarán estratificaciones más finas. En palabras simples, uniendo las estratificaciones para las elecciones locales es posible recuperar la estratificación por Distrito Federal. Para las elecciones locales los miembros del comité han usado distintas ideas para construir estratificaciones:

- Estratificar por municipio, en este caso se puede ver que uniendo municipios se obtienen los distritos federales.
- Intersección entre distritos federales y alguna otra variable. En este caso, uniendo todas las intersecciones de un mismo distrito se obtiene el distrito completo.
- Estratificar usando distritos locales siempre y cuando uniendo varios distritos locales se obtengan distritos federales.

El manejo de estratificaciones más finas en los estados tiene dos objetivos:

1. Alcanzar mayor precisión para la elección local con menor muestra.
2. Lograr una mayor dispersión de la muestra.

Otra diferencia fundamental es que en general los tamaños de muestra para las elecciones locales implicarán tamaños de muestra significativamente más grandes por distrito federal a los buscados en la elección de diputados. Entonces, en estados con elección para gobernador, la muestra para la elección federal será una sub-muestra de la elección local. Sin embargo, en el caso excepcional en el que en algún distrito federal la muestra para la elección local sea menor a la muestra de la elección federal, la muestra para la gubernatura se convertirá en una sub-muestra de la federal.

Bajo las consideraciones descritas anteriormente, es posible demostrar analíticamente que las probabilidades de selección de casillas para la elección federal son las mismas que se obtendrían asumiendo un diseño estratificado por distrito federal directamente. Por lo tanto, estos detalles no representan ningún problema técnico.

Intervalo consolidado

El Dr. Luis Enrique Nieto y el Maestro Carlos Pérez realizarán una estimación de la conformación de la cámara de diputados siguiendo una perspectiva Bayesiana de la inferencia estadística. En nuestro caso, se seguirá la alternativa clásica vía las ideas de re-muestreo Bootstrap descritas anteriormente. Para consolidar ambas estimaciones en una sola, se usará la unión. Por lo que se obtendrá una sola estimación de la conformación de la Cámara de Diputados que cubrirá la conformación real con al menos un 95% de confianza. Se seguirá el mismo procedimiento para consolidar las estimaciones para la participación y los porcentajes con respecto a la VVE.

Referencias

Efron, B. (1979). "Bootstrap methods: Another look at the jackknife". The Annals of Statistics, Vol 7. No 1, pp. 1-26.

5.1.2. Procedimiento de estimación

5.1.2.1. Dr. Luis Enrique Nieto Barajas¹⁴

Metodología

El problema de estimación de la conformación de la cámara, desde un punto de vista estadístico, requiere de hacer estimaciones a dos niveles: primero, a nivel de cada uno de los 300 distritos federales, para los diputados elegidos por mayoría simple; y en segundo lugar, a nivel nacional para la determinación de los 200 diputados por representación proporcional. Para el primer caso, se necesita tener representación muestral a nivel distrito federal, y para el segundo, se necesita tener representación a nivel nacional. Es por esto que el diseño muestral más apropiado para el problema en cuestión es un muestreo aleatorio estratificado por distrito federal.

El reglamento de elecciones en sus artículos 413 y 414 establece los principios para la determinación de los diputados electos por mayoría relativa, y el artículo 415 junto con sus anexos 420, 421 y 422 establecen los principios para la elección de diputados por representación proporcional.

Aunque los criterios se basan en la votación total y los votos por partido, es posible representar los criterios en términos de las proporciones de votos. Para este efecto introducimos la siguiente notación.

Sea θ_{ij} la proporción de votos, relativa al listado nominal, a favor del candidato/partido j en el estrato (distrito) i , con $j = 1, \dots, J$ e $i = 1, \dots, N$. J es el número total de candidatos (que incluye a los candidatos por partido político y los independientes) más dos categorías extras: los votos nulos y no registrados y las abstenciones. $N = 300$ es el número de distritos federales. Sea n_i el listado nominal de votantes en el distrito i para $i = 1, \dots, N$, y $n = \sum_{i=1}^N n_i$ el tamaño de la lista nominal nacional. Entonces $\theta_j = \sum_{i=1}^N \frac{n_i}{n} \theta_{ij}$ es la proporción de votos, relativa al listado nominal, a favor del candidato j a nivel nacional.

Para la determinación de los diputados ganadores por mayoría relativa definimos

$$mr_{ij} = \begin{cases} 1, & \text{si } \theta_{ij} > \theta_{il} \text{ } l \neq j, l = 1, \dots, J - 2 \\ 0, & \text{en otro caso} \end{cases}$$

¹⁴ Con la colaboración del Mtro. Carlos Samuel Pérez Pérez

una indicadora que nos dice si el candidato j ganó o no en el distrito i . Finalmente, definimos $mr_j = \sum_{i=1}^N mr_{ij}$ como el número de diputaciones ganadas por el partido j por el principio de mayoría relativa.

Para la determinación de diputados ganadores por representación proporcional definimos primero $\lambda_j = \frac{\theta_j}{\sum_{l=1}^{J-2} \theta_l}$ la proporción de votos válidos (se excluyen las abstenciones), y $\eta_j = \frac{\theta_j}{\sum_{l=1}^{J-3} \theta_{ll}(\lambda_j > 0.03)}$ como la proporción después de excluir a los partidos que no alcanzan el 3% en la votación válida, los independientes y los nulos y no registrados.

El algoritmo iterativo consiste en calcular primero $rp_j = \lfloor 200\eta_j \rfloor$ los curules por representación proporcional cruda para el partido j y se ajusta de acuerdo con la sobrerrepresentación, i.e.,

$$\text{si } mr_j + rp_j > \lfloor 500(\eta_j + 0.08) \rfloor \text{ entonces } rp_j^* = \lfloor 500(\eta_j + 0.08) \rfloor - mr_j.$$

Posteriormente se calculan los decimales remanentes $200\eta_j - rp_j$ y se ordenan de mayor a menor y se les asigna una diputación más a cada uno hasta completar 200.

Finalmente, el número de escaños por partido político NE_j resulta de sumar los escaños por mayoría relativa mr_j y los de representación proporcional rp_j^* . Es decir, $NE_j = mr_j + rp_j^*$.

Para la estimación de los parámetros antes descritos, y en particular para estimar θ_{ij} , usaremos un enfoque bayesiano. En específico usaremos el modelo propuesto por Mendoza y Nieto-Barajas (2016)¹⁵.

Dicho modelo supone que X_{ij}^k , denota los votos a favor del partido/candidato j en el estrato i en la casilla k . Esta variable observable sigue un comportamiento descrito por la distribución siguiente

$$X_{ij}^k | \theta_{ij}, \tau_{ij} \sim N \left(n_i^k \theta_{ij}, \frac{\tau_{ij}}{n_i^k} \right)$$

¹⁵ Mendoza, M. y Nieto-Barajas L.E. (2016). Quick counts in the Mexican presidential elections: A Bayesian approach. *Electoral Studies* 43, 124–132.

con $k = 1, \dots, c_i$, con c_i casillas en muestra en el distrito i , $i = 1, \dots, N$, y $j = 1, \dots, J$. Donde n_i^k es la lista nominal para la casilla k en el distrito i y τ_{ij} es la precisión que se supone constante dentro del estrato e independiente de θ_{ij} , más aún X_{ij}^k independiente $X_{ij}^{k'}$, para $j \neq j'$.

La verosimilitud para cada estrato $i = 1, \dots, N$ y cada candidato $j = 1, \dots, J$ está dada por

$$L(\theta_{ij}, \tau_{ij} | x_{ij}) \propto \tau_{ij}^{\frac{c_i}{2}} \exp \left\{ -\frac{\tau_{ij}}{2} \sum_{k=1}^{c_i} \frac{1}{n_i^k} (x_{ij}^k - n_i^k \theta_{ij})^2 \right\}$$

La distribución inicial para los parámetros desconocidos del modelo, θ_{ij} y τ_{ij} se considera no informativa y toma la forma

$$p(\theta_{ij}, \tau_{ij}) \propto \tau_{ij}^{-1} I(\tau_{ij} > 0) I(0 < \theta_{ij} < 1)$$

donde $I(A)$ denota la función indicadora para el conjunto A .

Así, usando el Teorema de Bayes, la distribución posterior resulta ser proporcional al producto de una normal truncada para θ_{ij} condicional en τ_{ij} y una distribución gamma para τ_{ij} . Los estimadores puntuales y por intervalo para θ_{ij} se aproximarán mediante simulaciones de la distribución posterior.

5.1.2.2. Dr. Carlos Erwin Rodríguez Hernández-Vela¹⁶

Estimación vía re-muestreo Bootstrap

Para realizar la estimación de la conformación de la Cámara de Diputados se utilizará la técnica estándar de re-muestreo Bootstrap. Es importante mencionar que dada la complejidad para obtener la conformación de la cámara de diputados no es posible aplicar fórmulas analíticas para realizar la estimación. Sin embargo, los métodos de re-muestreo proporcionan una herramienta, que como ya se vio en el Conteo Rápido de 2015, resuelve el problema de forma por demás satisfactoria.

A continuación, se brinda una descripción general del re-muestreo Bootstrap para estimar la conformación de la Cámara de Diputados:

De la muestra de casillas del primer distrito se seleccionará una sub-muestra de casillas y con ésta se estimará el número total de votos en favor de cada contendiente (candidatos con partido o independiente). Con esta información se pronosticará el contendiente ganador del primer distrito. Se replicará este procedimiento en cada uno de los 299 distritos restantes.

A continuación, se usará la información de las sub-muestras para estimar la VTE¹⁷, VVE¹⁸ y VNE¹⁹. Finalmente se aplicarán las reglas de repartición proporcional de escaños para obtener la configuración de los 200 diputados restantes (ver Apéndice).

Con este procedimiento se obtendrá una conformación completa de 500 diputados. Esto constituye una iteración de la técnica de re-muestreo conocida como Bootstrap, ver por ejemplo Efron (1979), y este proceso se repetirá tantas veces como sea posible. Se pretende realizar 10,000 iteraciones obteniendo así 10,000 configuraciones. Con este número de conformaciones se calculará el intervalo de confianza del 95% para el punto uno.

Las estimaciones de la participación nacional en la elección, así como del porcentaje de votación para cada partido con respecto a la VVE se pueden realizar utilizando técnicas estándar de estimación o vía la información de los re-

¹⁶ Con la colaboración del Act. Edgar Gerardo Alarcón González

¹⁷ VTE: Votación Total Emitida

¹⁸ VVE: Votación Válida Emitida

¹⁹ VNE: Votación Nacional Emitida

muestreos Bootstrap. Nosotros utilizaremos la información de los re-muestreos para hacer todo en el mismo procedimiento.

Observaciones:

- En cada distrito será imperativo contar con información de al menos 5 casillas de las seleccionadas en muestra. Esto es de la mayor importancia debido a que se requiere hacer inferencia en cada uno de los 300 distritos electorales: existe una correspondencia uno a uno entre los trescientos escaños por mayoría relativa y el resultado en cada uno de los 300 distritos.
- La estrategia de re-muestreo es computacionalmente demandante, el procesamiento en paralelo será una herramienta clave que nos ayudará a generar el mayor número posible de iteraciones en el menor tiempo posible.

Apéndice: Conformación de la Cámara de Diputados

En este apéndice se describe la forma de obtener la conformación de la Cámara de Diputados a partir de la votación observada en cada uno de los 300 distritos federales electorales. De acuerdo con el artículo 53 de la Constitución Política de los Estados Unidos Mexicanos, este procedimiento se divide en dos partes: 1) **Mayoría Relativa**, en donde se otorgan 300 escaños por cada uno de los distritos federales electorales, y 2) **Representación Proporcional**, en donde se otorgan otros 200 escaños para aquellos partidos que satisfagan los *requisitos* establecidos en la ley.

Mayoría Relativa

- Con base en el artículo 15 de la LEGIPE se obtiene la *Votación Total Emitida* (VTE), esta se obtiene sumando el total de votos depositados en las urnas. Para efectos de la implementación de Representación Proporcional, es importante observar que la VTE se puede descomponer por partidos, candidatos independientes, candidatos no registrados y votos nulos.
- Tomando como fundamento el mismo artículo, se obtiene la *Votación Válida Emitida* (VVE), esta se obtiene restando de la VTE los votos nulos y a candidatos no registrados. De acuerdo con el artículo 54 de la Constitución, para que un partido conserve su registro y tenga derecho a conseguir

escaños por Representación Proporcional será necesario que posea al menos el 3% de la composición de la VVE.

- Fundamentado en el artículo 15 de la LEGIPE, en aplicación del artículo 54 de la constitución, se define la *Votación Nacional Emitida* (VNE) la cual está conformada por aquellos partidos políticos (sin incluir candidatos independientes) que satisfacen el *requisito* antes mencionado.
- Finalmente, los escaños bajo el principio de Mayoría Relativa se otorgarán a los contendientes que obtengan el mayor número de *votos emitidos* en cada uno de los 300 Distritos Electorales Federales.

Nota: En caso de existir coaliciones, será necesario distribuir los *votos emitidos* entre los partidos que la conforman, considerando los acuerdos de coalición establecidos.

Representación Proporcional

Para realizar la asignación de los 200 escaños bajo este principio, haremos uso de algunas de las estadísticas obtenidas en el principio de Mayoría Relativa. Recordemos que este principio está sujeto a los partidos políticos que satisfacen lo establecido en la Constitución (*ver artículo 54*).

- Con fundamento en el artículo 16 de la LEGIPE, se calcula el *Cociente Natural* (CN) como la división de la VNE entre 200. Esto se puede interpretar como la cantidad de votos que vale cada escaño.
- Con base en la cantidad anterior y de acuerdo con los artículos 16 y 17 de la LEGIPE, se obtiene el mayor entero menor o igual a la división entre la *votación emitida* de cada partido y el CN para realizar un primer reparto los escaños. En caso de existir lugares sobrantes, estos se repartirán bajo el principio de *resto mayor* el cual consiste en ver el residuo de la división antes mencionada y asignar los escaños sobrantes equitativamente a los escaños más grandes, habiendo así otorgado en un primer reparto los escaños por Representación Proporcional.
- Para evitar la sobrerrepresentación, la LEGIPE en el artículo 17 establece condiciones límite para el total de escaños por ambos principios (Mayoría Relativa + Representación Proporcional), cantidad que no puede ser superada por ningún partido político y que en caso de que así sea, se le

deberán restar lugares otorgados por Representación Proporcional hasta satisfacer la cota mencionada. Dichos escaños penalizados se deberán reasignar entre los partidos que no hayan superado los límites establecidos (LEGIPE - artículo 18).

Este procedimiento debe repetirse hasta que la repartición cumpla las restricciones establecidas por la Ley.

Nota: Con fundamento en el artículo 17 de la LEGIPE, los límites para evitar la sobrerrepresentación penalizarán únicamente los escaños otorgados por Representación Proporcional sin afectar la cantidad de escaños obtenidos por Mayoría Relativa.

5.2. Elecciones de Gobernatura

5.2.1. Estratificación y tamaño de muestra

5.2.1.1. Equipo 1

5.2.1.1.1. Baja California Sur ²⁰

I. Elementos para diseño muestral

a. El parámetro de precisión

Los conteos rápidos constituyen un instrumento para que la ciudadanía y actores interesados puedan conocer en forma preliminar un primer estimado del resultado de la jornada electoral. Los resultados se expresan, no como estimaciones puntuales sino como intervalos en el porcentaje de votación, cada uno con una confianza determinada. Ciertamente no se sabe cuál será el sentido de voto de la población en la próxima Jornada Electoral y por ello, un criterio de relevancia en el diseño es el tratar de establecer un tamaño de muestra tal, que se espere que los intervalos de estimación sean reducidos y que por tanto, disminuya la probabilidad de que exista un traslape en la noche de la Jornada Electoral. Para hacer estos estudios se analizaron distintos esquemas de estratificación del estado y de tamaños de la muestra.

1. Escenarios de estratificación explorados

Para la realización de estos trabajos se utilizaron los resultados electorales en la última elección para gobernador del estado y de los resultados, restringidos a esta entidad federativa, en la elección de Presidente de la República 2018. Los parámetros básicos de estas elecciones se presentan a continuación.

²⁰ Elaborado por el Dr. Alberto Alonso y Coria con la colaboración del Fis. Joaquín Morales Bolio

Marco Geográfico Electoral del estado de Baja California Sur

Elección	Distritos		Casillas				Total casillas
	Federales	Locales	Urbanas	Mixtas	Rurales	Mixtas + Rurales	
Gobernador 2015	2	16	614	163	127	290	904
Presidente 2018	2	16	666	129	183	312	978

Tabla 1*

* Elaborado con base en los resultados electorales proporcionados por el INE

Primeramente, se exploraron las siguientes estratificaciones²¹:

- Distritos federales.
- Distritos Locales.
- Distritos Locales y Federales separando la parte rural y urbana. (explorándose como criterio de separación el 30% y 40%. Esto es, bifurcando el estrato sólo si el distrito contaba con al menos ese porcentaje de casillas rurales o no-urbanas).

Para cada uno de los escenarios y cada uno de los tamaños de muestra, se realizaron mil muestras distintas y se obtuvo el promedio de la precisión. El promedio que se reporta a continuación es la peor precisión (la más grande) para cada uno de los partidos o en su caso coaliciones que compitieron en la elección en cuestión. Los cortes del 30% o 40% sólo se usaron de manera indicativa para la posible exploración de un corte más fino. No todos los distritos locales o federales se separaban en subestratos urbano-mixto-rural (o urbano y rural-mixto) en razón de la inviabilidad práctica de la división por no haber suficientes casillas de alguno de los tipos que diera lugar a estratos suficientemente poblados en términos de número de casillas. Se anexa una tabla que sirvió de guía para determinar en qué distritos si se consideró viable la exploración rural-urbano (la tabla presentada es la correspondiente a la elección de gobernador. Una semejante se utilizó en el caso del análisis de la elección de presidente).

²¹ Se utilizó el marco cartográfico vigente en 2020

Baja California Sur												
Distrito / Estrato	Distritos Federales			Distritos Locales			Manual 1			Manual 2		
	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas
1	519	13.29%	25.24%	47	4.26%	89.36%	160	31.25%	53.13%	106	26.42%	52.83%
2	384	14.84%	38.02%	67			174	10.92%	20.11%	107	25.23%	38.32%
3				64		6.25%	185		5.95%	121	11.57%	19.01%
4				54		12.96%	224	25.45%	61.16%	185		5.95%
5				61	22.95%	37.70%	160		5.63%	117	47.01%	77.78%
6				64	51.56%	73.44%				107	1.87%	42.99%
7				60		6.67%				160		5.63%
8				49		18.37%						
9				52								
10				53	9.43%	22.64%						
11				54	40.74%	53.70%						
12				53	41.51%	83.02%						
13				62	27.42%	48.39%						
14				44	25.00%	59.09%						
15				60								
16				59								

Tabla 2*

* Elaborado con base en las casillas para la elección de gobernador 2015 con el marco cartográfico vigente. Marco proporcionado por el INE

Un punto de relevancia en el diseño fue la consideración de que una submuestra de la que se plantea utilizar para Baja California Sur pudiera ser utilizada para el ejercicio de conteo rápido para la estimación de la conformación de la Cámara de Diputados. Así pues, se consideró pertinente que los estratos estuvieran contenidos en los distritos electorales federales en razón de que estos serán utilizados para ese ejercicio.

En una segunda etapa se exploraron distintas agrupaciones de distritos locales. Las dos más notables se describen a continuación:

Baja California Sur					
Agrupación 1			Agrupación 2		
DF	DL	ESTRATO	DF	DL	ESTRATO
1	11, 13, 14	1	1	13, 14	1
	5, 10, 15	2		10, 11	2
	2, 3, 4	3		5, 15	3
2	1, 6, 7, 12	4		2, 3, 4	4
	8, 9, 16	5		6, 12	5
			2	1, 7	6
				8, 9, 16	7

Tabla 3

2. Precisión vs. Tamaño de muestra.

La siguiente tabla presenta el análisis correspondiente a la elección de gobernador 2015.

Baja California Sur. Elección Gobernador														
No. Estratos	2	3	16	19	20	19	5	6	7	6	7	8	10	9
	Distritos Federales		Distritos Locales			Agrupación 1				Agrupación 2				
Tamaño de la muestra	Distritos Federales	con subestratos mixtos 30%	Distritos Locales	con subestratos rurales 30%	con subestratos mixtos 30%	con subestratos mixtos 40%	Agrupación 1	con subestratos rurales 30%	con subestratos mixtos 30%	con subestratos mixtos 40%	Agrupación 2	con subestratos rurales 30%	con subestratos mixtos 30%	con subestratos mixtos 40%
60	1.90	1.93	1.63	1.57	1.59	1.61	1.90	1.89	1.87	1.88	1.80	1.82	1.81	1.75
70	1.74	1.76	1.53	1.51	1.58	1.49	1.72	1.69	1.66	1.70	1.72	1.71	1.66	1.69
80	1.64	1.62	1.43	1.42	1.39	1.39	1.63	1.65	1.57	1.56	1.58	1.61	1.50	1.59
90	1.57	1.53	1.37	1.37	1.33	1.35	1.52	1.51	1.51	1.55	1.42	1.46	1.43	1.43
100	1.48	1.48	1.26	1.27	1.23	1.24	1.43	1.38	1.42	1.42	1.39	1.37	1.36	1.37
110	1.38	1.39	1.22	1.16	1.18	1.19	1.35	1.37	1.32	1.36	1.32	1.33	1.29	1.27
120	1.32	1.32	1.14	1.15	1.16	1.15	1.32	1.28	1.28	1.29	1.24	1.24	1.24	1.22
130	1.28	1.26	1.11	1.11	1.06	1.07	1.22	1.26	1.22	1.24	1.20	1.22	1.18	1.18
140	1.22	1.20	1.05	1.06	1.04	1.04	1.16	1.18	1.20	1.18	1.18	1.12	1.13	1.11
150	1.17	1.18	0.99	1.02	1.01	0.98	1.14	1.13	1.12	1.13	1.09	1.11	1.08	1.07
160	1.10	1.13	0.98	0.98	0.96	0.98	1.10	1.09	1.08	1.10	1.06	1.04	1.02	1.02
170	1.08	1.09	0.93	0.93	0.93	0.94	1.06	1.05	1.04	1.05	1.03	1.01	1.00	1.00
180	1.04	1.05	0.89	0.89	0.89	0.88	1.01	1.01	1.00	0.99	1.00	1.00	0.96	0.98
190	1.01	1.00	0.88	0.86	0.86	0.86	0.98	0.98	0.97	0.99	0.96	0.94	0.91	0.93
200	0.98	0.97	0.85	0.83	0.84	0.83	0.96	0.95	0.94	0.94	0.93	0.91	0.90	0.90

Tabla 4. precisión elección gobernador 2015

Como puede observarse la estratificación con distritos locales con una muestra de 150 casillas genera una precisión inferior al uno por ciento.

En un ejercicio semejante realizado con los datos de la elección presidencial 2018, si bien coincidentemente apuntan a la bondad de la estratificación vía distritos electorales locales, no es sino hasta un tamaño de muestra de 180 casillas que se obtendría una precisión inferior al uno por ciento. (Tabla 5).

BAJA CALIFORNIA SUR. Elección federal																
No. Estratos	2	3	16	18	17	20	5	6	5	7	6	7	9	8	10	9
	Distritos Federales		Distritos Locales				Agrupación 1				Agrupación 2					
Tamaño de la muestra	Distritos Federales	con subestratos mixtos 30%	Distritos Locales	con subestratos rurales 30%	con subestratos rurales 40%	con subestratos mixtos 30%	Agrupación 1	con subestratos rurales 30%	con subestratos rurales 40%	con subestratos mixtos 30%	con subestratos mixtos 40%	Agrupación 2	con subestratos rurales 30%	con subestratos rurales 40%	con subestratos mixtos 30%	con subestratos mixtos 40%
110	1.55	1.52	1.36	1.36	1.36	1.39	1.51	1.53	1.51	1.52	1.50	1.53	1.49	1.48	1.49	1.49
120	1.46	1.43	1.27	1.28	1.28	1.27	1.43	1.47	1.43	1.45	1.46	1.44	1.40	1.39	1.39	1.42
130	1.41	1.36	1.24	1.24	1.24	1.22	1.38	1.39	1.38	1.39	1.41	1.37	1.34	1.34	1.34	1.33
140	1.33	1.33	1.18	1.16	1.17	1.17	1.34	1.51	1.34	1.33	1.31	1.32	1.27	1.27	1.29	1.30
150	1.31	1.27	1.12	1.10	1.09	1.15	1.29	1.43	1.29	1.25	1.28	1.28	1.23	1.23	1.21	1.22
160	1.25	1.23	1.08	1.06	1.09	1.09	1.24	1.38	1.24	1.21	1.24	1.22	1.18	1.19	1.18	1.19
170	1.20	1.18	1.05	1.04	1.03	1.07	1.17	1.34	1.17	1.16	1.18	1.17	1.16	1.15	1.14	1.15
180	1.17	1.15	1.01	1.00	0.99	1.01	1.15	1.29	1.15	1.13	1.13	1.13	1.09	1.11	1.10	1.10
190	1.13	1.12	0.98	0.98	0.98	0.95	1.10	1.24	1.10	1.10	1.11	1.10	1.04	1.06	1.06	1.07
200	1.09	1.06	0.95	0.95	0.95	0.96	1.07	1.17	1.07	1.07	1.06	1.06	1.04	1.04	1.04	1.04

Tabla 5. Precisión elección federal 2018

3. Número de casillas por estrato

Un parámetro a considerar es el número de casillas muestra por estrato. Éste es relevante, en tanto que es posible que por vicisitudes del día de la jornada

electoral pudiera no ser posible la obtención de la información de alguna de ellas. Es por ello que es de importancia tener en cuenta la siguiente tabla:

Baja California Sur															
Tamaño de la muestra	Distritos Federales		Distritos Locales				Agrupación 1			Agrupación 2					
	Distritos Federales	con substratos mixtos 30%	Distritos Locales	con substratos rurales 30%	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 1	con substratos rurales 30%	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 2	con substratos rurales 30%	con substratos mixtos 30%	con substratos mixtos 40%	
60	26 - 34	10 - 34	3 - 4	2 - 4	2 - 4	2 - 4	11 - 15	3 - 15	5 - 12	5 - 15	7 - 12	4 - 12	3 - 12	3 - 12	
70	30 - 40	11 - 40	3 - 5	2 - 5	2 - 5	2 - 5	12 - 17	4 - 17	6 - 14	6 - 17	8 - 14	4 - 14	3 - 14	4 - 14	
80	34 - 46	13 - 46	4 - 6	2 - 6	2 - 6	2 - 6	14 - 20	4 - 20	7 - 16	7 - 20	9 - 16	5 - 16	4 - 16	4 - 16	
90	38 - 52	15 - 52	4 - 7	2 - 7	2 - 7	2 - 7	16 - 22	5 - 22	7 - 18	7 - 22	11 - 18	5 - 18	4 - 18	5 - 18	
100	43 - 57	16 - 57	5 - 7	2 - 7	2 - 7	2 - 7	18 - 25	6 - 25	8 - 20	8 - 25	12 - 20	6 - 20	5 - 20	5 - 20	
110	47 - 63	18 - 63	5 - 8	3 - 8	2 - 8	2 - 8	19 - 27	6 - 27	9 - 23	9 - 27	13 - 23	7 - 23	5 - 23	6 - 23	
120	51 - 69	19 - 69	6 - 9	3 - 9	2 - 9	2 - 9	21 - 30	7 - 30	10 - 25	10 - 30	14 - 25	7 - 25	5 - 25	6 - 25	
130	55 - 75	21 - 75	6 - 10	3 - 10	3 - 10	3 - 10	23 - 32	7 - 32	11 - 27	11 - 32	15 - 27	8 - 27	6 - 27	7 - 27	
140	60 - 80	23 - 80	7 - 10	3 - 10	3 - 10	3 - 10	25 - 35	8 - 35	12 - 29	12 - 35	16 - 29	9 - 29	6 - 29	7 - 29	
150	64 - 86	24 - 86	7 - 11	4 - 11	3 - 11	3 - 11	27 - 37	8 - 37	12 - 31	12 - 37	18 - 31	9 - 31	7 - 31	8 - 31	
160	68 - 92	26 - 92	8 - 12	4 - 12	3 - 12	3 - 12	28 - 40	9 - 40	13 - 33	13 - 40	19 - 33	10 - 33	7 - 33	8 - 33	
170	72 - 98	27 - 98	8 - 13	4 - 13	3 - 13	3 - 13	30 - 42	9 - 42	14 - 35	14 - 42	20 - 35	10 - 35	8 - 35	9 - 35	
180	77 - 103	29 - 103	9 - 13	4 - 13	4 - 13	4 - 13	32 - 45	10 - 45	15 - 37	15 - 45	21 - 37	11 - 37	8 - 37	9 - 37	
190	81 - 109	31 - 109	9 - 14	5 - 14	4 - 14	4 - 14	34 - 47	11 - 47	16 - 39	16 - 47	22 - 39	12 - 39	9 - 39	10 - 39	
200	85 - 115	32 - 115	10 - 15	5 - 15	4 - 15	4 - 15	35 - 50	11 - 50	17 - 41	17 - 50	23 - 41	12 - 41	9 - 41	10 - 41	

Tabla 6.*

* Datos de la elección gobernador 2015

b. La situación con los Capacitadores Asistentes Electorales (CAE)

Una limitante para el ejercicio del conteo rápido es la necesidad de un intenso trabajo en campo durante la noche del día de celebración de los comicios. En razón de que la muestra de casillas se genera en cada estrato en forma aleatoria, es posible que algunos CAE tengan la responsabilidad de reportar más de una casilla. Así pues, un parámetro de relevancia consiste en el estimar el porcentaje de CAE que tendrán sólo una casilla para reportar y que, por tanto, puedan acudir en forma oportuna a su ubicación y proceder, conforme a los lineamientos, a informar los resultados. Quizá sea importante mencionar que en estudios realizados por organización electoral²² no existe una relación totalmente directa en la oportunidad de reporte y el número de casillas a atender por el CAE. Sin embargo, es claro que se facilita la operación en campo si se reduce el número de casillas a atender por esta figura y por ende su estudio formó parte de los parámetros a considerar.

²² 10feb2021_DEOE_PRINCIPALES RESULTADOS CR 2018 y 2019.pptx. De hecho, sólo el 70% de los CAE con una sola casilla reportaron a tiempo para el conteo rápido federal.

Baja California Sur	
Porcentaje de CAEs con una casilla de la muestra	
Tamaño de la muestra	Distritos Locales
60	93.01%
70	90.87%
80	88.78%
90	87.10%
100	85.13%
110	83.59%
120	81.99%
130	80.11%
140	78.72%
150	76.48%
160	74.98%
170	73.44%
180	71.67%
190	70.40%
200	68.56%

Tabla 7*

*Datos elección gubernamental

c. Arribo de la muestra en el caso de la elección presidencial 2018

A diferencia del PREP que por su naturaleza en los primeros minutos u horas tiene un posible sesgo geográfico producido por la distinta velocidad de llegada de las actas a las sedes distritales entre, digamos, las casillas urbanas y rurales, el conteo rápido cuenta con un medio más eficiente de comunicación que minimiza o disminuye esta problemática. En razón de la posible existencia de una comunicación menos eficiente en las casillas rurales, se estudió el posible efecto sobre la hora de llegada de la información de las casillas en la muestra. Se realizó un estudio sobre la forma de arribo de la información de las casillas muestra utilizando los datos correspondientes a Baja California Sur del conteo rápido celebrado para la elección presidencial. Los datos y gráficas asociadas se presentan a continuación.

Datos BCS 2018				
	Urbanas	Mixtas	Rurales	Mixtas + Rurales
Total de casillas ¹	666	129	183	312
	68.10%	13.19%	18.71%	31.90%
Casillas marco muestral ³	31	11	5	16
	65.96%	23.40%	10.64%	34.04%
Casillas capturadas ²	29	10	3	13
	69.05%	23.81%	7.14%	30.95%

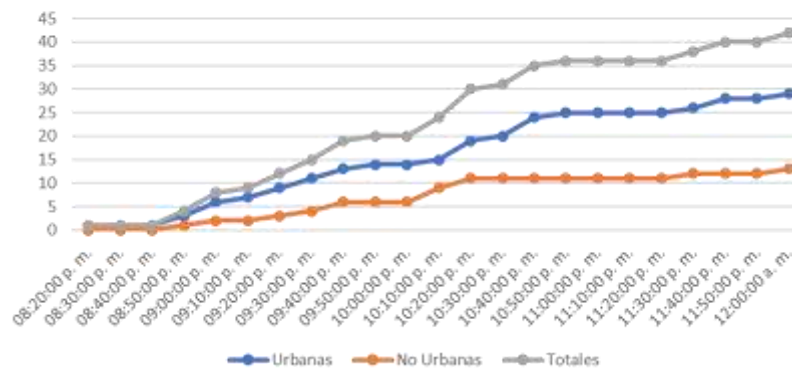
Tabla 8

¹ Listado_casillas_2018.csv

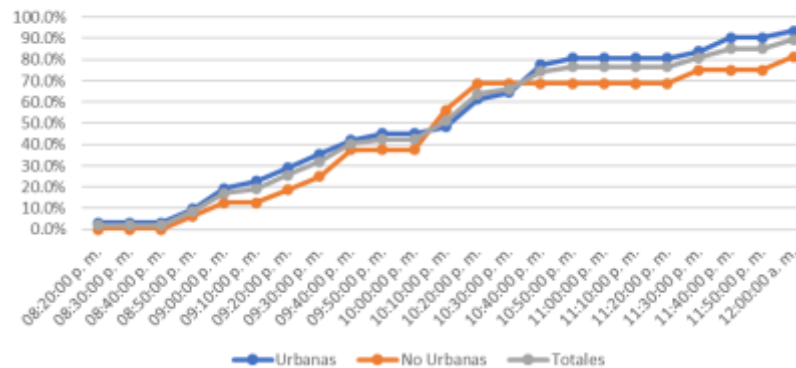
² Corte TOTALES0100020040-txt 2 jul 2018 12:40 a.m.

³ 4-ConteoRapido18MUESTRA-ELECCION-PRESIDENCIAL.csv

Acumulado por tipo de sección
BCS 2018



Porcentaje de llegada por tipo de sección
BCS 2018



Acorde a la gráfica derecha, es de destacar que Baja California Sur es un estado que en el conteo rápido presidencial 2018 presentó sólo una diferencia pequeña en la velocidad de captura entre casillas urbanas y no-urbanas. Esto ciertamente abona a no incluir en la estratificación posibles subestratos.

d. Estratificación recomendada

Del análisis de los datos obtenidos se concluye que la estratificación por distritos locales es un buen esquema muestral, no solo permite el esperar una buena precisión, sino que garantiza además una adecuada cobertura geográfica. La tabla resumen siguiente nos indica que el tamaño de muestra de 130 casillas sería el adecuado para realizar la estimación.

Baja California Sur			
No Estratos = 16			
	Precisión (%)		
	Elección Gobernador		Elección federal
Tamaño de la muestra	Criterio Razón	Criterio remuestreo local*	Criterio Razón
120	1.14	1.17	1.27
130	1.11	1.10	1.24
140	1.05	1.06	1.17
150	0.99	1.05	1.11
160	0.98	0.98	1.08

Tabla 9

* Datos proporcionados por Dr. Gabriel Nuñez y M. en C. Emiliano Geneyro

e. Expansión por arriba.

Un parámetro a considerar es el arribo de la muestra. En el ejercicio realizado en el proceso electoral 2017-18 en las entidades que tuvieron elección de gobernador se pudo observar que de la muestra solicitada arribó un porcentaje bajo. Como puede observarse en la tabla siguiente el mayor porcentaje de

muestra recibida fue el de Morelos con 80.0% y el peor fue el de Tabasco con un 58.2%. ²³

Conteos rápidos elecciones Gobernador 2018				
Entidad	Porcentaje recibido	Casillas totales	Casillas en muestra	Muestra recibida
Cdmx	58.4	12,976	1,108	647
Chiapas	59.2	6,316	500	294
Guanajuato	71.4	7,463	500	357
Jalisco	73.7	9,820	472	348
Morelos	80	2,421	200	160
Puebla	73.1	7,547	509	372
Tabasco	58.4	2,912	450	263
Veracruz	70.1	10,595	1,100	771
Yucatán	62.7	2,666	300	188
Total		62,716	5,139	3,400

Tabla 10*

* informe final COTECORA

Así pues, se considera en el diseño, que sólo arribará cerca del 67% de la muestra y por tanto se establece como parámetro de una muestra efectiva $130/.67 \sim 190^{24}$ casillas. Si bien la fracción muestral es alta, esto es derivado de que es la segunda entidad más pequeña del País en término de número de casillas. El tamaño permite a su vez estimar que más del 70% de los CAE tendrán que reportar sólo una casilla.

f. Tabla final de Estratos

Para establecer la tabla final de estratos se utilizó el listado de casillas correspondiente al estado de Baja California Sur, aprobadas por los consejos distritales con corte al 30 de marzo de 2021. Se utilizarán los distritos locales como estratificación.

Estratificación: Distritos locales.

Precisión estimada: 1.1% a 1.24%

²³ Informe COTECORA junio 2018.

²⁴ Redondeado a la decena más próxima.

Baja California Sur					
ESTRATO	Estado	Distrito federal	Distrito Local	N Casillas estrato (Nh)	N Casillas en muestra (nh)
1	3	2	1	65	12
2	3	1	2	66	12
3	3	1	3	65	12
4	3	1	4	77	14
5	3	1	5	71	13
6	3	2	6	70	13
7	3	2	7	68	13
8	3	2	8	64	12
9	3	2	9	67	12
10	3	1	10	56	10
11	3	1	11	59	11
12	3	2	12	62	11
13	3	1	13	65	12
14	3	1	14	48	10
15	3	1	15	61	11
16	3	2	16	66	12
			TOTAL	1030	190

Tabla 11

II. Algunas consideraciones

1. Se debe señalar que la precisión en las estimaciones se maneja tanto en la etapa de planeación del ejercicio como en la presentación de los resultados finales del mismo. En la etapa de planeación se usa para determinar un tamaño de muestra adecuado teóricamente, para alcanzar el nivel de precisión y confianza deseadas en las estimaciones. Sin embargo, debido a que se fija este tamaño de muestra tomando como referencia los resultados de elecciones previas y a que se tendrá un porcentaje de no respuesta (real) que no conocemos hasta el día de la jornada electoral, una vez concluida la elección y con toda la información disponible se determinará la precisión realmente obtenida.
2. Los resultados de las simulaciones presentadas son únicamente para el método de estimación usado, y tienen como objetivo fijar un tamaño de

muestra. Sin embargo, se pueden emplear métodos de estimación alternativos que arrojarán precisiones distintas.

3. Se debe señalar que hasta el momento (marzo de 2021) se cuenta con la proyección de 259 CAE (información proporcionada por la DEOE), de los cuales se ha estimado que en promedio 148 apoyarían el CR para la elección a la Gobernatura en el Estado de Baja California Sur. De esta estimación de CAE, el 70.1% (103) solo tendrá que reportar 1 sola casilla. Por lo cual, es pertinente señalar que el tamaño de muestra de 190 casillas se encuentra dentro de los límites de la capacidad operativa de campo.
4. Hay que mencionar, por un lado, que si se recibe un tamaño de muestra menor a 190 casillas, el día de la Jornada Electoral, todavía se pueden tener estimaciones confiables si la distribución de la muestra recibida es adecuada. Sin embargo, las precisiones serán menores como se puede ver en la Tabla 9. Por ejemplo, aún para un tamaño de muestra de 130 casillas con una composición y distribución suficiente desde el punto de vista teórico y que corresponda al Diseño muestral aquí definido, se alcanzaría una precisión alrededor de 1.1%, pero se podría tener un sesgo importante en las estimaciones reportadas. Por otro lado, también se debe señalar que si se recibe una muestra menor a 130 casillas con una composición y calidad muy alejada de la aquí definida por el Diseño muestral, difícilmente se podrán obtener resultados que ofrezcan certidumbre en las estimaciones.
5. Finalmente hay que señalar que los diferentes procedimientos de estimación que implementarán los tres miembros del COTECORA del grupo 1 (Mtra. Claudia Ortiz, Dr. Alberto Alonso y Dr. Gabriel Núñez), para la realización del CR del Estado de Baja California Sur, estarán basados en el Diseño muestral definido en esta Sección.

III. Integración de las estimaciones

Cada uno de los 3 miembros del grupo 1 del COTECORA, con la misma información de las casillas recibidas el día de la jornada electoral, realizarán una estimación para la votación de cada uno de los candidatos a la Gobernatura del Estado de Baja California Sur, así como, para el porcentaje de participación. Por lo anterior, para integrar los resultados anteriores se propone usar la unión de al menos dos de las tres estimaciones obtenidas. Sin embargo, la manera

definitiva de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros.

5.2.1.1.2. Guerrero²⁵

Diseño muestral para el Conteo Rápido de la elección a la Gobernatura

El objetivo general de los ejercicios de Conteo Rápido (CR) es contribuir a dar certidumbre, confianza y transparencia en un proceso electoral. Lo anterior, ofreciendo una estimación de las tendencias de la votación el mismo día de la jornada electoral. Se debe enfatizar que el éxito de los Conteo Rápidos depende del trabajo conjunto de varios actores: Los miembros del Comité Técnico Asesor del Conteo Rápido (COTECORA), el INE, los miembros de las Juntas Locales, los miembros de los Organismos Públicos Locales Electorales (OPLE), etc.

Desde el punto de vista estadístico un Conteo Rápido es un procedimiento que con base en la información de una muestra aleatoria de casillas de escrutinio y cómputo, permite estimar con cierta precisión y confianza, las tendencias de los resultados la misma noche de la elección.

Uno de los aspectos técnicos responsabilidad del COTECORA es el denominado Diseño Muestral, el cual se puede definir como el conjunto de estrategias y procedimientos definidos para seleccionar una muestra de una población objetivo de estudio, la cual debe cumplir con ciertas características estadísticas adecuadas. El Diseño muestral también considera procedimientos que minimicen los errores de sesgo y cobertura, asociados a los procedimientos de estimación que se emplean. Así, desde un punto de vista moderno, se puede decir que un Diseño muestral queda definido por la determinación de un tamaño de muestra, una estrategia de selección de la muestra y un método de estimación. Lo anterior, con la finalidad de lograr niveles aceptables de precisión en las estimaciones ofrecidas, con un alto nivel de confianza estadística.

A continuación, se presentan algunos análisis y consideraciones técnicas que permitieron definir el Diseño muestral que se empleará para el Conteo Rápido en la elección del 6 de junio de 2021 a la Gobernatura del Estado de Guerrero.

²⁵ Elaborado por el Dr. Gabriel Núñez Antonio con la colaboración del Mtro. Emiliano Geneyro Squarzón

Estrategia de Selección

Considerando el Artículo 373 del Reglamento de Elecciones (donde se establece que la muestra debe abarcar la mayor dispersión geográfica posible) y tomando en cuenta la experiencia técnica adquirida en anteriores ejercicios de Conteo Rápido, se concluye que la estrategia de selección de las casillas de escrutinio y cómputo sea un muestreo aleatorio estratificado. Un aspecto importante en un muestreo estratificado es la definición o construcción de los estratos, los cuales son sub-grupos de la población que cumplen con dos características: no se traslapan y la unión de todos los estratos conforma a toda la población objetivo.

Para el caso del Conteo Rápido del Estado de Guerrero, nuestra población está conformada por el total de casillas aprobadas para el día de la Jornada Electoral. Para definir los estratos se consideraron diferentes estratificaciones (forma de definir estratos) entre las cuales se encuentran:

- Estratificación por Distritos Federales.
- Estratificación considerando las intersecciones de Distritos Federales y Distritos Locales.
- Estratificación considerando las intersecciones por Distritos Federales y tipo de sección.
- Estratificación por Distritos Locales.
- Estratificación considerando las intersecciones Distritos Locales y tipo de sección.

Para tomar una decisión final de cuál es la mejor estratificación a emplear, se debe considerar el impacto de varios factores como:

1. La precisión (límite máximo deseable para el error de estimación) que se pretende obtener en las estimaciones reportadas.
2. Porcentaje de CAE que participarán reportando 1 sola casilla para el Conteo Rápido.
3. El tamaño de la muestra que se va a seleccionar.
4. El tamaño de muestra asignado en cada uno de los estratos.

5. El nivel de confianza deseado en las estimaciones, el cual se ha fijado en 95%.

Se debe mencionar que, para la definición del tamaño de muestra, adicionalmente se deben considerar factores como:

- El Proceso de llegada y la calidad de la información muestral (características de la muestra recibida y la hora de llegada).
- El Tamaño de la no-respuesta en ejercicios de Conteo Rápido anteriores.

Por otro lado, también se debe mencionar que en la definición/construcción del Diseño muestral se ha tomado en cuenta que alrededor del 80% de los CAE apoyando al Conteo Rápido, solo reporte 1 casilla. Lo anterior, ha sido una consideración que los miembros del COTECORA han asumido en la definición técnica del diseño muestral con el fin de contribuir a que la presión de los CAE el día de la jornada electoral no sea excesiva y, con esto propiciar condiciones (desde el punto de vista técnico-estadístico) para lograr la recepción de un mayor porcentaje de muestra el día de la Jornada Electoral.

A continuación, se presentan algunos conceptos probabilísticos para ir complementando algunos de los puntos antes mencionados.

Algunas ideas básicas del muestreo probabilístico para el Conteo Rápido

Uno de los objetivos técnicos del Conteo Rápido es estimar la proporción p de votos a favor de un candidato a la Gubernatura²⁶ del Estado de Guerrero. Para lo anterior, se selecciona una muestra aleatoria (bajo algún esquema de muestreo o estrategia de selección) de n casillas de un total de N y con la información muestral disponible se calcula un estimador \hat{p} . Mediante muestreo probabilístico es posible analizar el impacto de diferentes *estrategias de selección, tamaños de muestra n y estimadores*, para asegurar (teóricamente) que:

$$|p - \hat{p}| \leq d, \quad (1)$$

²⁶ Esto se hará para cada contendiente, así como, para la estimación de la votación en favor de candidatos no registrados y votación nula; sin embargo, para efectos de la explicación nos concentraremos en la estimación para un solo contendiente.

con un 95% de confianza. A la cantidad ***d*** comúnmente se le conoce como precisión o límite máximo aceptable para el error de estimación. La expresión (1) se puede escribir de forma equivalente en términos de un intervalo de confianza al $(1-\alpha) \times 100\%$, como

$$\hat{p} - d \leq p \leq \hat{p} + d \quad (2)$$

Esto significa, por ejemplo, que si $(1-\alpha)=0.95$, $d=0.02$, y se extraen 1,000 muestras distintas e independientes una de otra (cada una siguiendo la misma estrategia de selección, usando el mismo tamaño de muestra y el mismo estimador) y con cada muestra se hace una estimación, entonces se tendrían 1000 estimaciones puntuales $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4, \hat{p}_5, \dots, \hat{p}_{1000}$, de las cuales se esperaría que el 95% de estas estimaciones cumpliera con la condición (1). Es decir, que su distancia máxima (error de estimación) respecto al verdadero valor sea de 2 puntos porcentuales (2%) con una confianza del 95%.

Precisión en las estimaciones

El procedimiento anterior permite analizar diferentes diseños muestrales para obtener las precisiones bajo varias estrategias de selección, diferentes tamaños de muestra y diferentes estimadores. En consecuencia, poder decidir cuál es el diseño de muestreo más adecuado para llevar a cabo el procedimiento correspondiente del Conteo Rápido en la elección para la Gobernatura del Estado de Guerrero.

A continuación, se presentan y analizan las precisiones para diferentes estratificaciones y diferentes tamaños de muestra. En todos los análisis realizados se tomó como referencia los cómputos distritales de la elección a la Gobernatura del Estado de Guerrero en 2015 y los cómputos distritales de la elección presidencial de 2018, asociados al estado de Guerrero. Las estratificaciones consideradas en la definición del diseño maestro fueron:

1. Distritos federales \Rightarrow 9 estratos
2. Distritos federales y distritos locales \Rightarrow 32 estratos
3. Distritos federales y tipo de sección \Rightarrow 18 estratos
4. Distritos locales \Rightarrow 28 estratos
5. Distritos locales y tipo sección \Rightarrow 53 estratos

Para el análisis de las precisiones, se consideró el estimador de razón

$$\hat{p}_i = \frac{\sum_{h=1}^L N_h \bar{y}_{hi}}{\sum_i \sum_{h=1}^L N_h \bar{y}_{hi}} \quad (3)$$

En donde el sub-índice i representa al candidato i -ésimo (incluyendo también el número de votos por candidatos no registrados y nulos), h es el número de estrato, L el número de estratos, N_h es el número de casillas instaladas en el estrato h y \bar{y}_{hi} es el número promedio de votos en las casillas en la muestra para el candidato i en el estrato h . Las estimaciones se realizaron para todos los candidatos (incluyendo candidatos no registrados y nulos). Sin embargo, en aras de una mejor exposición sólo se muestra el análisis relativo al candidato que registró la mayor varianza, y por tanto requería un tamaño de muestra mayor, para alcanzar al menos la misma precisión en todos los candidatos.

La Tabla G.1 muestra los resultados promedio, basados en un estudio de 5000 simulaciones, de las precisiones alcanzadas para las diferentes estratificaciones y diferentes tamaños de muestra.

Tabla G.1. Análisis de precisiones para el CR a la Gubernatura de Guerrero

n	Dto. Fed ²⁷ (9 estratos)	Dto. Fed X Dto. Loc ²⁸ (32 estratos)	Dto. Fed. X Tipo Sección (18 estratos)	Dto. Loc (28 estratos)	Dto. Loc X Tipo Sección (53 estratos)
600	1.15	1.00	1.10	1.05	1.02
650	1.11	0.98	1.07	0.99	0.98
700	1.04	0.95	1.02	0.97	0.95
750	0.99	0.90	0.98	0.93	0.91
800	0.96	0.89	0.95	0.89	0.90
850	0.94	0.84	0.94	0.88	0.86
900	0.92	0.82	0.91	0.84	0.83

Se puede observar que las estrategias de selección de distritos locales por tipo de sección y distritos locales con intersección de distritos federales arrojan las mejores precisiones. Sin embargo, bajo la estratificación distritos locales por tipo de sección se tienen 53 estratos. Por la experiencia en otros ejercicios de Conteo

²⁷ Dto. Fed: Distrito Federal

²⁸ Dto. Loc: Distrito Local

Rápido, este número de estratos se consideran demasiados, ya que puede traer problemas en la proporción de muestra asignada en cada uno de estos estratos. Por otra parte, se puede observar que la estratificación de distrito federal por distrito local y la estratificación solo por distrito local (28 estratos) son muy parecidas en cuanto a las precisiones obtenidas para los diferentes tamaños de muestra analizados. Hay que señalar que, básicamente, estas dos últimas estratificaciones tienen como base a los distritos locales.

Presión sobre los CAE apoyando el CR

Por otro lado, dada la inquietud en la carga que han tenido los CAE en los últimos ejercicios de Conteo Rápido, se analizó también la proporción de CAE a los que les tocaría reportar 1, 2 y 3 o más casillas. Para este análisis se consideró la información de la elección a la Gobernatura en 2015, así como, la información de las ARE que se tenía en 2020. Los resultados se muestran en la Tabla G.2.

Tabla G.2. Porcentajes de CAE con 1, 2, y 3 o más casillas

n	Casillas x CAE	Dto. Fed (9 estratos)	Dto. Fed X Dto. Loc (32 estratos)	Dto. Fed. X Tipo Sección (18 estratos)	Dto. Loc (28 estratos)	Dto. Loc X Tipo Sección (53 estratos)
800	1	78.4%	79.2%	80.3%	79.2%	81.3%
	2	19.2%	18.7%	17.7%	18.7%	16.9%
	3 o más	2.4%	2.1%	2.0%	2.1%	1.8%
850	1	77.1%	77.9%	79.2%	77.8%	80.1%
	2	20.2%	19.7%	18.6%	19.8%	17.9%
	3 o más	2.7%	2.4%	2.2%	2.4%	2.0%
900	1	75.7%	76.6%	77.8%	76.5%	78.8%
	2	21.2%	20.7%	19.6%	20.8%	19.0%
	3 o más	3.1%	2.7%	2.6%	2.7%	2.2%

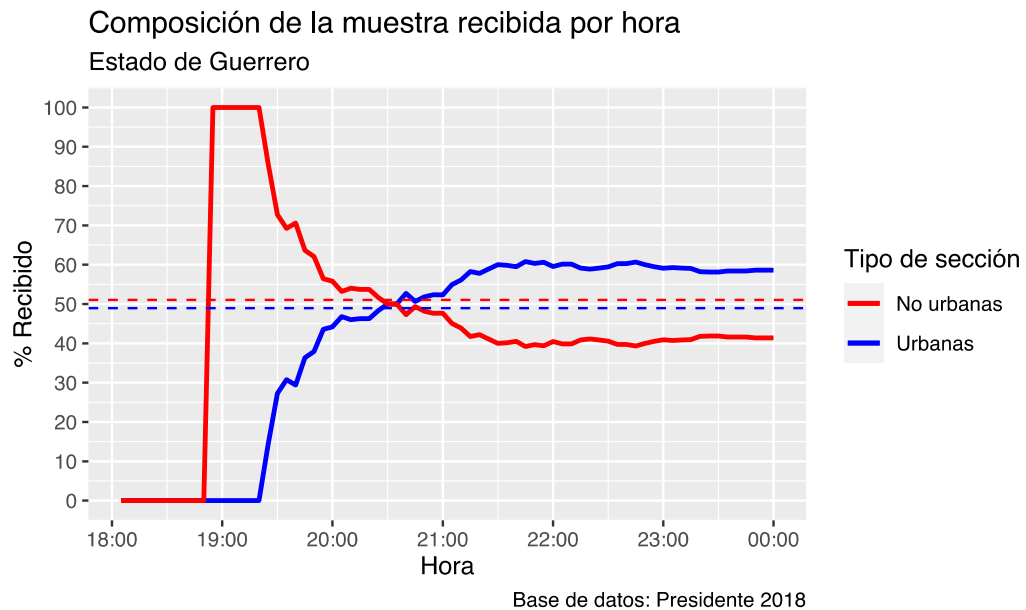
Analizando los resultados derivados de la Tabla G.2, se puede ver que si se consideran tamaños de muestra de 800 o 850 aún se tienen porcentajes altos (alrededor del 77.9%) de CAE a los cuales solo les tocaría reportar 1 casilla, bajo la estratificación de Distrito Federal por Distrito Local.

El Proceso de llegada y la calidad de la información muestral

Un aspecto importante en la determinación de la estrategia de selección es la calidad de la información recibida. En este sentido, se analizó la composición por tipo de sección de la llegada de la información de las casillas en muestra.

En el diseño muestral establecido en el COTECORA de 2018, la composición de casillas en muestra era de 49% Urbanas y 51% No-Urbanas. Sin embargo, como se puede ver en la Gráfica G.3, a las 11:00 pm del día de la jornada electoral, se había recibido una muestra con una composición contraria a la establecida en el diseño original. Es decir, a las 11:00 pm se contaba con una muestra con una composición del 59% de casillas Urbanas y 41% de casillas No-Urbanas.

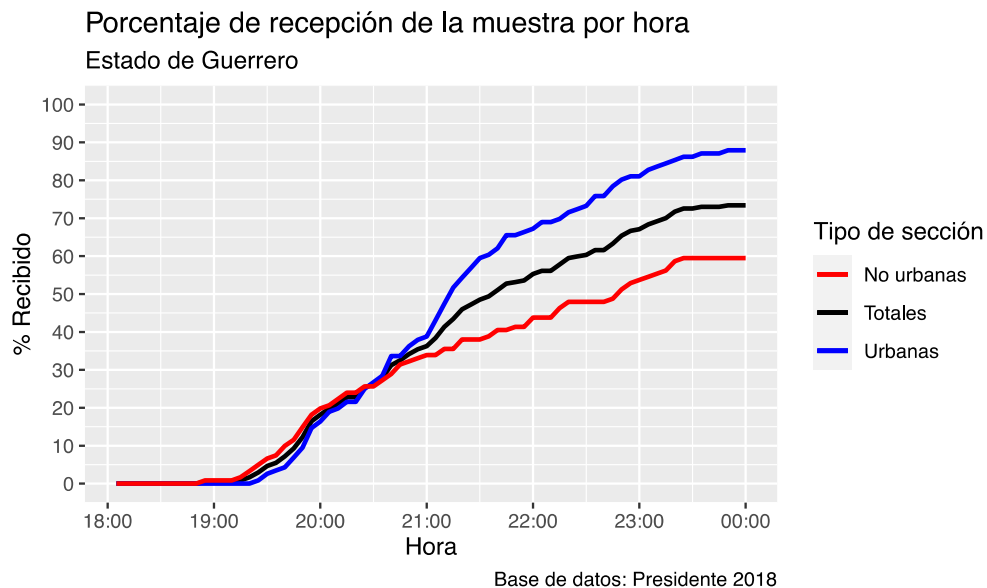
Gráfica G.3. Composición por Tipo de Sección de la muestra recibida.



Este aspecto de desbalance se puede presentar en las remesas que el día de la jornada electoral se reciban y se debe considerar en la determinación del Diseño muestral. En este caso, se hicieron análisis tomando en cuenta estos desbalances y se observó que aun teniendo estos desbalances en la composición de la muestra, la estratificación por distritos locales y el método de estimación empleado permiten subsanar este aspecto y siguen produciendo estimaciones razonables.

Como se mencionó anteriormente, otro aspecto importante en la determinación del tamaño de muestra es el proceso de arribo de la información de las casillas en muestra, el día de la jornada electoral. Para tener una perspectiva de este aspecto, se analizó la hora de llegada de la muestra relativa al Estado de Guerrero en el CR de la elección presidencial de 2018. La Gráfica G.4 muestra este proceso de llegada, diferenciando el proceso para casillas en secciones urbanas y no urbanas, así como, para el caso general.

Gráfica G.4. Proceso de llega (en porcentaje) de las casillas solicitadas en muestra



En 2018, el número de casillas en muestra en el estado de Guerrero fue de 237. De estas casillas a las 10:00 pm (hora del centro) del día de la jornada electoral, solo se habían recibido el 55.2% de la muestra solicitada. A las 11:00 pm se tenía el 67% y a las 12:00 am, solo se tenía poco más del 74% de la muestra solicitada. Este análisis es relevante ya que muestra la preocupación del flujo de llegada de la información para el COTECORA. Si esta tasa de llegada de la información se presenta en la jornada electoral del 6 de junio de 2021, difícilmente se podrán dar estimaciones con una precisión adecuada antes de las 11:00 pm.

Determinación del tamaño de muestra

Derivado del análisis anterior sobre el proceso de llegada, el día de la jornada electoral estaríamos esperando recibir alrededor de las 11:00 pm el 67% de la

muestra solicitada. Adicionalmente, derivado del análisis de la presión de los CAE que participaran el día de la jornada electoral, podemos decir que si elegimos un tamaño de muestra de 850 casillas al menos el 78% de ellos tendrían que reportar solo 1 casilla.

Derivado de todos los análisis presentados anteriormente, el tamaño de muestra solicitado para el CR a la Gobernatura del Estado de Guerrero, que nos permitiría tener estimaciones con precisiones adecuadas, a una hora de llegada razonable y sin imponer una carga excesiva a los CAE, es de $n=850$ casillas.

Estratificación final

Tomando en cuenta que la muestra para la elección federal (CR para la estimación de la conformación de la Cámara de Diputados) se hará basándose en distritos federales, para el CR de la elección a la Gobernatura de Guerrero se define una estratificación por distritos locales bajo las siguientes consideraciones:

- Separar los distritos locales que pertenezcan a diferentes distritos federales.
- Asignación de la muestra de tamaño $n=850$ en cada estrato (nh), de manera proporcional al número de casillas totales en cada estrato (Nh)
- Para evitar poca muestra en algún estrato y derivado de los resultados de un análisis de cobertura realizado, basado en simulaciones de 100 muestras de tamaño 600, se colapsaron (unieron) algunos distritos locales en un mismo estrato.

Al corte del 30 de marzo de 2021 se contaba con 5,007 casillas asignadas al Estado de Guerrero. Así, la estratificación final, tomando en cuenta todas las condiciones anteriores, se muestran en la Tabla G.5. Se debe mencionar que si hay un ajuste en el número de casillas los valores de Nh y nh presentarán modificaciones y esto hay que tenerlo muy presente, sobre todo para el sistema elaborado por el área de Cartografía de la DERFE.

Consideramos que este diseño muestral ofrece las condiciones técnicas necesarias para obtener resultados adecuados el día de la jornada electoral del 6 de junio de 2021 en el Conteo Rápido a la Gobernatura del Estado de Guerrero. Hay que mencionar que este tipo de estrategias de colapso ha sido empleado por el Dr. Gabriel Núñez Antonio en otros CR como en el de Veracruz en 2018, lo

cual ha permitido obtener estimaciones adecuadas a un menor costo de tamaño de muestra.

Tabla G.5 Definición de la Estratificación final para el CR a la Gobernatura del Estado de Guerrero. Total de casillas en cada estrato (Nh) y tamaños de muestra en cada estrato (nh). Las Nh y nh son calculadas con base al corte de casillas al 30 de marzo de 2021.

ESTRATO	DISTRITOS FEDERALES	DISTRITOS LOCALES	Nh	nh
1	1	17	203	34
2	1	18	220	37
3	1	20	217	37
4	2	21	200	34
5	2	22	174	30
6	2	23	191	32
7	3	8 y 10	285	48
8	3	11	185	31
9	3	12	182	31
10	4	3	151	26
11	4	4	169	29
12	4	5 y 8	208	35
13	5	15 y 16	140	24
14	5	27	189	32
15	5	28	204	35
16	6	24	195	33
17	6	25	190	32
18	6	26	205	35
19	7	1	155	26
20	7	2	174	30
21	7	19	160	27
22	8	13	173	29
23	8	14	170	29
24	8	15 y 16	186	32
25	9	3 y 7	190	32
26	9	6	140	24
27	9	9	151	26
		Totales:	5007	850

Con la estratificación anterior y usando nuevamente el estimador de razón definido en (3) se obtuvo una precisión de al menos 1.1 puntos porcentuales en las estimaciones, con un porcentaje de 78% de CAE con una sola casilla asignada.

Algunas consideraciones importantes

1. Se debe señalar que la precisión en las estimaciones se maneja tanto en la etapa de planeación del ejercicio como en la presentación de los resultados finales del mismo. En la etapa de planeación se usa para determinar un tamaño de muestra adecuado teóricamente, para alcanzar el nivel de precisión y confianza deseadas en las estimaciones. Sin embargo, debido a que se fija este tamaño de muestra tomando como referencia los resultados de elecciones previas y a que se tendrá un porcentaje de no respuesta (real) que no conocemos hasta el día de la jornada electoral, una vez concluida la elección y con toda la información disponible se determinará la precisión realmente obtenida.
2. Los resultados de las simulaciones presentadas son únicamente para el método de estimación usado, y tienen como objetivo fijar un tamaño de muestra. Sin embargo, se pueden emplear métodos de estimación alternativos que arrojarán precisiones distintas.
3. Se debe señalar que hasta el momento (marzo de 2021) se cuenta con la proyección de 1,324 CAE (información proporcionada por la DEOE), de los cuales se ha estimado que en promedio 640 apoyarían el CR para la elección a la Gubernatura en el Estado de Guerrero. De esta estimación de CAE, el 78% (500) solo tendrá que reportar 1 sola casilla. Por lo cual, es pertinente señalar que el tamaño de muestra de 850 casillas se encuentra dentro de los límites de la capacidad operativa de campo.
4. Hay que mencionar, por un lado, que si se recibe un tamaño de muestra menor a 850 casillas, el día de la Jornada Electoral, todavía se pueden tener estimaciones confiables si la distribución de la muestra recibida es adecuada. Sin embargo, las precisiones serán menores como se puede ver en la Tabla G.1. Por ejemplo, aún para un tamaño de muestra de 650 casillas con una composición y distribución suficiente desde el punto de vista teórico y que corresponda al Diseño muestral aquí definido, se alcanzaría una precisión alrededor de 1.2%, pero se podría tener un sesgo importante en las estimaciones reportadas. Por otro lado, también se debe

señalar que si se recibe una muestra menor a 650 casillas con una composición y calidad muy alejada de la aquí definida por el Diseño muestral, difícilmente se podrán obtener resultados que ofrezcan certidumbre en las estimaciones.

5. Finalmente hay que señalar que los diferentes procedimientos de estimación que implementarán los tres miembros del COTECORA del grupo 1 (Mtra. Claudia Ortiz, Dr. Alberto Alonso y Dr. Gabriel Núñez), para la realización del CR del Estado de Guerrero, estarán basados en el Diseño muestral definido en esta Sección.

Integración de las estimaciones

Cada uno de los 3 miembros del grupo 1 del COTECORA, con la misma información de las casillas recibidas el día de la jornada electoral, realizarán una estimación para la votación de cada uno de los candidatos a la Gubernatura del Estado de Guerrero, así como, para el porcentaje de participación. Por lo anterior, para integrar los resultados anteriores se propone usar la unión de al menos dos de las tres estimaciones obtenidas. Sin embargo, la manera definitiva de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros.

5.2.1.1.3. Nuevo León²⁹

Diseño muestral para el Conteo Rápido de la elección a la Gobernatura

El objetivo general de los ejercicios de Conteo Rápido (CR) es contribuir a dar certidumbre, confianza y transparencia en un proceso electoral. Lo anterior, ofreciendo una estimación de las tendencias de la votación el mismo día de la jornada electoral. Se debe enfatizar que el éxito de los Conteos Rápidos depende del trabajo conjunto de varios actores: Los miembros del Comité Técnico Asesor del CR (COTECORA), el INE, los miembros de las Juntas Locales, los miembros de los Organismos Públicos Locales Electorales (OPLE), etc.

Desde el punto de vista estadístico un Conteo Rápido es un procedimiento que con base en la información de una muestra aleatoria de casillas de escrutinio y cómputo, permite estimar con cierta precisión y confianza, las tendencias de los resultados la misma noche de la elección.

Uno de los aspectos técnicos de responsabilidad del COTECORA es el denominado Diseño Muestral, el cual se puede definir como el conjunto de estrategias y procedimientos definidos para seleccionar una muestra de una población objetivo de estudio, la cual debe cumplir con ciertas características estadísticas adecuadas. El Diseño muestral también considera procedimientos que minimicen los errores de sesgo y cobertura, asociados a los procedimientos de estimación que se emplean. Así, desde un punto de vista moderno, se puede decir que un Diseño muestral queda definido por la determinación de un tamaño de muestra, una estrategia de selección de la muestra y un método de estimación. Lo anterior, con la finalidad de lograr niveles aceptables de precisión en las estimaciones ofrecidas, con un alto nivel de confianza estadística.

A continuación, se presentan algunos análisis y consideraciones técnicas que permitieron definir el Diseño muestral que se empleará para el Conteo Rápido en la elección del 6 de junio de 2021 a la Gobernatura del Estado de Nuevo León.

²⁹ Elaborado por el Dr. Gabriel Núñez Antonio con la colaboración del Mtro. Emiliano Geneyro Squarzón

Estrategia de Selección

Considerando el Artículo 373 del Reglamento de Elecciones (donde se establece que la muestra debe abarcar la mayor dispersión geográfica posible) y tomando en cuenta la experiencia técnica adquirida en anteriores ejercicios de Conteo Rápido, se concluye que la estrategia de selección de las casillas de escrutinio y cómputo sea un muestreo aleatorio estratificado. Un aspecto importante en un muestreo estratificado es la definición o construcción de los estratos, los cuales son sub-grupos de la población que cumplen con dos características: no se traslapan y la unión de todos los estratos conforma a toda la población objetivo.

Para el caso del Conteo Rápido del Estado de Nuevo León, nuestra población está conformada por el total de casillas aprobadas para el día de la Jornada Electoral. Para definir los estratos se consideraron diferentes estratificaciones (forma de definir estratos) entre las cuales se encuentran:

- Estratificación por Distritos Federales.
- Estratificación considerando las intersecciones de Distritos Federales y Distritos Locales.
- Estratificación considerando las intersecciones por Distritos Federales y tipo de sección.
- Estratificación por Distritos Locales.
- Estratificación considerando las intersecciones Distritos Locales y tipo de sección.

Para tomar una decisión final de cuál es la mejor estratificación a emplear, se debe considerar el impacto de varios factores como:

1. La precisión (límite máximo deseable para el error de estimación) que se pretende obtener en las estimaciones reportadas.
2. Porcentaje de CAE que participarán reportando 1 sola casilla para el Conteo Rápido.
3. El tamaño de la muestra que se va a seleccionar.
4. El tamaño de muestra asignado en cada uno de los estratos.

5. El nivel de confianza deseado en las estimaciones, el cual se ha fijado en 95%.

Se debe mencionar que, para la definición del tamaño de muestra, adicionalmente se deben considerar factores como:

- El Proceso de llegada y la calidad de la información muestral (características de la muestra recibida y la hora de llegada).
- El Tamaño de la no-respuesta en ejercicios de Conteo Rápido anteriores.

Por otro lado, también se debe mencionar que en la definición/construcción del Diseño muestral se ha tomado en cuenta que alrededor del 80% de los CAE apoyando al Conteo Rápido solo reporte 1 casilla. Lo anterior, ha sido una consideración que los miembros del COTECORA han asumido en la definición técnica del Diseño muestral con el fin de contribuir a que la presión de los CAE el día de la jornada electoral no sea excesiva y, con esto propiciar condiciones (desde el punto de vista técnico-estadístico) para lograr la recepción de un mayor porcentaje de muestra el día de la jornada electoral.

A continuación, se presentan algunos conceptos probabilísticos para ir complementando algunos de los puntos antes mencionados.

Algunas ideas básicas del muestreo probabilístico para el Conteo Rápido

Uno de los objetivos técnicos del Conteo Rápido es estimar la proporción p de votos a favor de un candidato a la Gubernatura³⁰ del Estado de Nuevo León. Para lo anterior, se selecciona una muestra aleatoria (bajo algún esquema de muestreo o estrategia de selección) de n casillas de un total de N y con la información muestral disponible se calcula un estimador \hat{p} . Mediante muestreo probabilístico es posible analizar el impacto de diferentes *estrategias de selección, tamaños de muestra n y estimadores*, para asegurar (teóricamente) que:

$$|p - \hat{p}| \leq d, \quad (1)$$

³⁰ Esto se hará para cada contendiente, así como, para la estimación de la votación en favor de candidatos no registrados y votación nula; sin embargo, para efectos de la explicación nos concentraremos en la estimación para un solo contendiente.

con un 95% de confianza. A la cantidad ***d*** comúnmente se le conoce como precisión o límite máximo aceptable para el error de estimación. La expresión (1) se puede escribir de forma equivalente en términos de un intervalo de confianza al $(1-\alpha) \times 100\%$, como

$$\hat{p} - d \leq p \leq \hat{p} + d \quad (2)$$

Esto significa, por ejemplo, que si $(1-\alpha)=0.95$, $d=0.02$, y se extraen 1,000 muestras distintas e independientes una de otra (cada una siguiendo la misma estrategia de selección, usando el mismo tamaño de muestra y el mismo estimador) y con cada muestra se hace una estimación, entonces se tendrían 1000 estimaciones puntuales $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4, \hat{p}_5, \dots, \hat{p}_{1000}$, de las cuales se esperaría que el 95% de estas estimaciones cumpliera con la condición (1). Es decir, que su distancia máxima (error de estimación) respecto al verdadero valor sea de 2 puntos porcentuales (2%) con una confianza del 95%.

Precisión en las estimaciones

El procedimiento anterior permite analizar diferentes diseños muestrales para obtener las precisiones bajo varias estrategias de selección, diferentes tamaños de muestra y diferentes estimadores. En consecuencia, poder decidir cuál es el diseño de muestreo más adecuado para llevar a cabo el procedimiento correspondiente del Conteo Rápido en la elección para la Gobernatura del Estado de Nuevo León.

A continuación, se presentan y analizan las precisiones para diferentes estratificaciones y diferentes tamaños de muestra. En todos los análisis realizados se tomó como referencia los cómputos distritales de la elección a la Gobernatura del Estado de Nuevo León en 2015 y los cómputos distritales de la elección presidencial de 2018, asociados al estado de Nuevo León. Las estratificaciones consideradas en la definición del diseño muestral fueron:

1. Distritos federales \Rightarrow 12 estratos
2. Distritos federales y distritos locales \Rightarrow 38 estratos
3. Distritos federales y tipo de sección \Rightarrow 19 estratos
4. Distritos locales \Rightarrow 26 estratos
5. Distritos locales y tipo sección \Rightarrow 36 estratos

Para el análisis de las precisiones, se consideró el estimador de razón

$$\hat{p}_i = \frac{\sum_{h=1}^L N_h \bar{y}_{hi}}{\sum_i \sum_{h=1}^L N_h \bar{y}_{hi}} \quad (3)$$

En donde el sub-índice i representa al candidato i -ésimo (incluyendo también el número de votos por candidatos no registrados y nulos), h es el número de estrato, L el número de estratos, N_h es el número de casillas instaladas en el estrato h y \bar{y}_{hi} es el número promedio de votos en las casillas en la muestra para el candidato i en el estrato h . Las estimaciones se realizaron para todos los candidatos (incluyendo candidatos no registrados y nulos). Sin embargo, en aras de una mejor exposición sólo se muestra el análisis relativo al candidato que registró la mayor varianza, y por tanto requería un tamaño de muestra mayor, para alcanzar al menos la misma precisión en todos los candidatos.

La Tabla NL.1 muestra los resultados promedio, basados en un estudio de 5000 simulaciones, de las precisiones alcanzadas para las diferentes estratificaciones y diferentes tamaños de muestra.

Tabla NL.1. Análisis de precisiones para el CR a la Gobernatura de Nuevo León

n	Dto. Fed (12 estratos)	Dto. Fed X Dto. Loc (38 estratos)	Dto. Fed. X Tipo Sección (19 estratos)	Dto. Loc (26 estratos)	Dto. Loc X Tipo Sección (36 estratos)
600	0.85	0.78	0.80	0.80	0.78
650	0.82	0.76	0.78	0.77	0.73
700	0.78	0.74	0.74	0.75	0.70
750	0.78	0.70	0.73	0.73	0.69
800	0.75	0.68	0.70	0.68	0.67
850	0.72	0.65	0.70	0.67	0.64
900	0.70	0.64	0.66	0.64	0.63

Se puede observar que las estrategias de selección de distritos locales por tipo de sección y distritos locales con intersección de distritos federales arrojan las mejores precisiones. Para estar en concordancia, por un lado, con el Diseño muestral de CR para la Conformación de la Cámara de Diputados en el cual se considerarán los Distritos Federales como estratos, y por otro lado, dado que (como veremos más adelante) el Tipo de Sección no aporta mayores beneficios en el diseño, se considera inicialmente una estratificación de Distrito Federal por Distrito Local. Bajo esta estratificación se tienen 38 estratos y por la experiencia en otros

ejercicios de Conteo Rápido, este número de estratos se consideran demasiados, ya que puede traer problemas en la proporción de muestra asignada en cada uno de estos estratos. Es decir, se puede tener una muestra más pulverizada (muy pequeña) en cada estrato. Hay que señalar que, básicamente, esta estratificación tiene como base a los distritos locales, por lo que para evitar el problema de tener un tamaño de muestra pulverizado, se usará un procedimiento de colapsar (unir) distritos locales en un mismo estrato.

Presión sobre los CAE apoyando el CR

Por otro lado, dada la inquietud en la carga que han tenido los CAE en los últimos ejercicios de Conteo Rápido, se analizó también la proporción de CAE a los que les tocaría reportar 1, 2 y 3 o más casillas. Para este análisis se consideró la información de la elección a la Gubernatura en 2015, así como, la información de las ARE que se tenía en 2020. Los resultados se muestran en la Tabla NL.2.

Tabla NL.2. Porcentajes de CAE con 1, 2, y 3 o más casillas

n	Casillas por CAE	Dto. Fed (12 estratos)	Dto. Fed X Dto. Loc (38 estratos)	Dto. Fed. X Tipo Sección (19 estratos)	Dto. Loc (26 estratos)	Dto. Loc X Tipo Sección (36 estratos)
800	1	79.8%	80.4%	80.6%	80.3%	81.2%
	2	18.0%	17.6%	17.3%	17.6%	16.9%
	3 o más	2.2%	2.0%	2.1%	2.1%	1.9%
850	1	78.6%	79.2%	79.4%	79.1%	80.0%
	2	18.9%	18.6%	18.3%	18.6%	17.9%
	3 o más	2.5%	2.2%	2.3%	2.3%	2.1%
900	1	77.3%	77.9%	78.2%	77.8%	78.8%
	2	19.9%	19.6%	19.2%	19.6%	18.8%
	3 o más	2.8%	2.5%	2.6%	2.6%	2.4%

Analizando los resultados derivados de la Tabla NL.2, se puede ver que si se consideran tamaños de muestra de 800 o 850 aún se tienen porcentajes altos (hasta del 79.2%) de CAE a los cuales solo les tocaría reportar 1 casilla, bajo una estratificación por Distrito Local con Distrito Federal.

El Proceso de llegada y la calidad de la información muestral

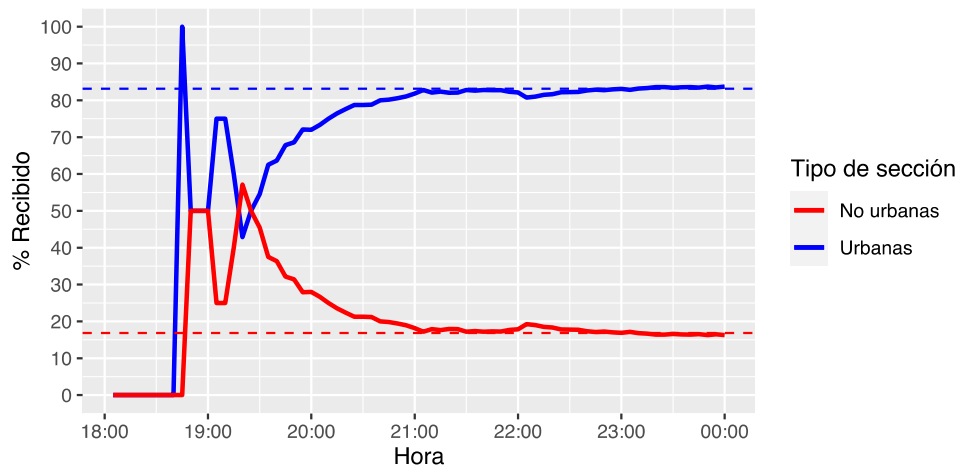
Un aspecto importante en la determinación de la estrategia de selección es la calidad de la información recibida. En este sentido, se analizó la composición por tipo de sección de la llegada de la información de las casillas en muestra.

En el diseño muestral establecido en el COTECORA de 2018, la composición de casillas en muestra era de 83.2% Urbanas y 16.8% No-Urbanas. Como se puede ver en la Gráfica NL.3, a las 11:00 pm del día de la jornada electoral, se había recibido una muestra con una composición acorde a la establecida en el diseño original. Es decir, a las 11:00 pm se contaba con una muestra con una composición del 83% de casillas Urbanas y 17% de casillas No-Urbanas. Dado este análisis se pudo concluir que el proceso de llegada de la información (remesas) no está influenciado o presenta algún desbalance en la composición de la muestra por el tipo de sección al que pertenece una casilla (de la muestra).

Gráfica NL.3. Composición por Tipo de Sección de la muestra recibida.

Composición de la muestra recibida por hora

Estado de Nuevo León

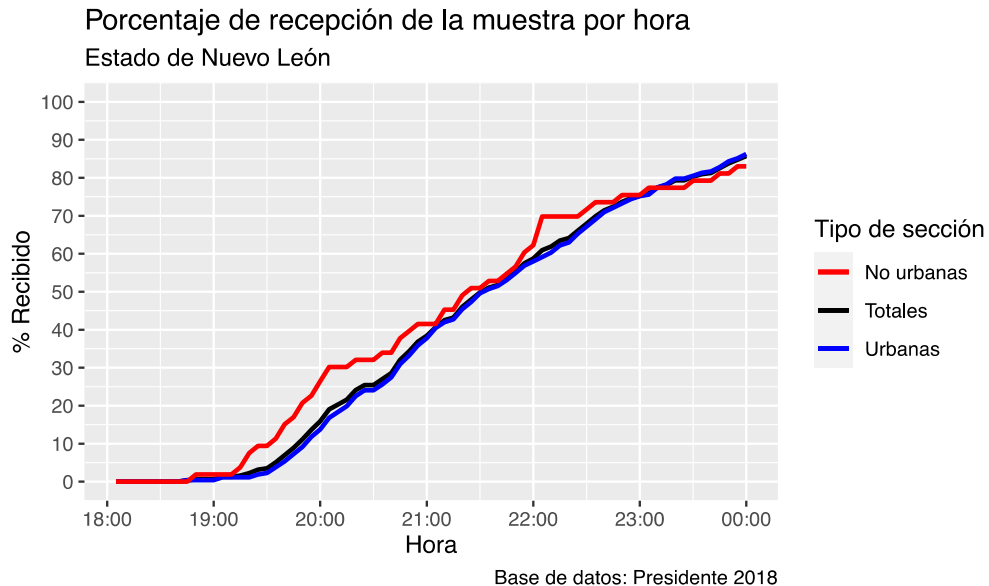


Base de datos: Presidente 2018

Como se mencionó anteriormente, otro aspecto importante en la determinación del tamaño de muestra es el proceso de arribo de la información de las casillas en muestra, el día de la jornada electoral. Para tener una perspectiva de este aspecto, se analizó la hora de llegada de la muestra relativa al Estado de Nuevo León en el CR de la elección presidencial de 2018. La Gráfica NL.4 muestra este

proceso de llegada, diferenciando el proceso para casillas en secciones urbanas y no urbanas, así como, para el caso general.

Gráfica N.L.4. Proceso de llega (en porcentaje) de las casillas solicitadas en muestra



En 2018, el número de casillas en muestra en el estado de Nuevo León fue de 315. De estas casillas a las 10:00 pm (hora del centro) del día de la jornada electoral, solo se habían recibido el 58.7% de la muestra solicitada. A las 10:35 pm se tenía el 68%. A las 11:00 pm se tenía el 75.2% y a las 12:00 am, se había alcanzado el 85.7% de la muestra solicitada. Este análisis es relevante ya que aunque muestra un flujo aceptable de llegada de la información para el COTECORA, se esperaría que en la jornada del 6 de junio de 2021 se tuviera una mayor recepción de muestra, para poder salir con estimaciones con una precisión adecuada en una hora más temprana a las 11:00 pm.

Determinación del tamaño de muestra

Derivado del análisis anterior sobre el proceso de llegada, el día de la jornada electoral estaríamos esperando recibir alrededor de las 10:35 pm el 68% de la muestra solicitada. Adicionalmente, derivado del análisis de la presión de los CAE que participaran el día de la jornada electoral, podemos decir que si elegimos un tamaño de muestra de 850 casillas al menos el 79% de ellos tendrían que reportar solo 1 casilla.

Derivado de todos los análisis presentados anteriormente, el tamaño de muestra solicitado para el CR a la Gubernatura del Estado de Nuevo León, que nos permitiría tener estimaciones con precisiones adecuadas, a una hora de llegada razonable y sin imponer una carga excesiva a los CAE, es de $n=850$ casillas.

Estratificación final

Tomando en cuenta que la muestra para la elección federal (CR para la estimación de la conformación de la Cámara de Diputados) se hará basándose en distritos federales, para el CR de la elección a la Gubernatura de Nuevo León se define una estratificación por distritos locales bajo las siguientes consideraciones:

- Separar los distritos locales que pertenezcan a diferentes distritos federales.
- Asignación de la muestra de tamaño $n=850$ en cada estrato (nh), de manera proporcional al número de casillas totales en cada estrato (Nh)
- Para evitar poca muestra en algún estrato y derivado de los resultados de un análisis de cobertura realizado, basado en simulaciones de 100 muestras de tamaño 600, se colapsaron (unieron) algunos distritos locales en un mismo estrato.

Al corte del 30 de marzo de 2021 se contaba con 6,971 casillas asignadas al Estado de Nuevo León. Así, la estratificación final, tomando en cuenta todas las condiciones anteriores, se muestran en la Tabla NL.5. Se debe mencionar que si hay un ajuste en el número de casillas los valores de Nh y nh presentarán modificaciones y esto hay que tenerlo muy presente, sobre todo para el sistema elaborado por el área de Cartografía de la DERFE.

Consideramos que este diseño muestral ofrece las condiciones técnicas necesarias para obtener resultados adecuados el día de la jornada electoral del 6 de junio de 2021 en el Conteo Rápido a la Gubernatura del Estado de Nuevo León. Hay que mencionar que este tipo de estrategias de colapso ha sido empleado por el Dr. Gabriel Núñez Antonio en otros CR como en el de Veracruz en 2018, lo cual ha permitido obtener estimaciones adecuadas a un menor costo de tamaño de muestra.

Tabla NL.5 Definición de la Estratificación final para el CR a la Gobernatura del Estado de Nuevo León. Total de casillas en cada estrato (Nh) y tamaños de muestra en cada estrato (nh). Las Nh y nh son calculadas con base al corte de casillas al 30 de marzo de 2021.

ESTRATO	DISTRITOS FEDERALES	DISTRITOS LOCALES	Nh	nh
1	1	18	265	32
2	1	19	256	31
3	2	5 y 7	378	46
4	2	16	223	27
5	3	17	240	29
6	3	25	264	32
7	4	9	243	30
8	4	10 y 11	373	46
9	5	1	182	22
10	5	2 y 11	220	27
11	6	1,2 y 3	331	41
12	6	4,6 y 11	294	36
13	7	20	413	50
14	7	21 y 24	330	40
15	8	7,13 y 15	304	37
16	8	14	208	26
17	9	23 y 26	320	39
18	9	24	312	38
19	10	4 y 8	288	35
20	10	6 y 11	322	39
21	11	12	231	28
22	11	13 y 15	314	38
23	12	22	278	34
24	12	23	382	47
		Totales:	6971	850

Con la estratificación anterior y usando nuevamente el estimador de razón definido en (3) se obtuvo una precisión de al menos 1.0 punto porcentual en las estimaciones, con un porcentaje de 79% de CAE con una sola casilla asignada.

Algunas consideraciones importantes

1. Se debe señalar que la precisión en las estimaciones se maneja tanto en la etapa de planeación del ejercicio como en la presentación de los resultados finales del mismo. En la etapa de planeación se usa para determinar un tamaño de muestra adecuado teóricamente, para alcanzar el nivel de precisión y confianza deseadas en las estimaciones. Sin embargo, debido a que se fija este tamaño de muestra tomando como referencia los resultados de elecciones previas y a que se tendrá un porcentaje de no respuesta (real) que no conocemos hasta el día de la jornada electoral, una vez concluida la elección y con toda la información disponible se determinará la precisión realmente obtenida.
2. Los resultados de las simulaciones presentadas son únicamente para el método de estimación usado, y tienen como objetivo fijar un tamaño de muestra. Sin embargo, se pueden emplear métodos de estimación alternativos que arrojarán precisiones distintas.
3. Se debe señalar que hasta el momento (marzo de 2021) se cuenta con la proyección de 1,727 CAE (información proporcionada por la DEOE), de los cuales se ha estimado que en promedio 660 apoyarían el CR para la elección a la Gubernatura en el Estado de Nuevo León. De esta estimación de CAE el 79% (522) solo tendrá que reportar 1 sola casilla. Por lo cual, es pertinente señalar que el tamaño de muestra de 850 casillas se encuentra dentro de los límites de la capacidad operativa de campo.
4. Hay que mencionar, por un lado, que si se recibe un tamaño de muestra menor a 850 casillas, el día de la Jornada Electoral, todavía se pueden tener estimaciones confiables si la distribución de la muestra recibida es adecuada. Sin embargo, las precisiones serán menores como se puede ver en la Tabla NL.1. Por ejemplo, aún para un tamaño de muestra de 650 casillas con una composición y distribución suficiente desde el punto de vista teórico y que corresponda al Diseño muestral aquí definido, se alcanzaría una precisión alrededor de 1.0%, pero se podría tener un sesgo importante en las estimaciones reportadas. Por otro lado, también se debe señalar que si se recibe una muestra menor a 650 casillas con una composición y calidad muy alejada de la aquí definida por el Diseño muestral, difícilmente se podrán obtener resultados que ofrezcan certidumbre en las estimaciones.
5. Finalmente hay que señalar que los diferentes procedimientos de estimación que implementarán los tres miembros del COTECORA del grupo

1 (Mtra. Claudia Ortiz, Dr. Alberto Alonso y Dr. Gabriel Núñez), para la realización del CR del Estado de Nuevo León, estarán basados en el Diseño muestral definido en esta Sección.

Integración de las estimaciones

Cada uno de los 3 miembros del grupo 1 del COTECORA, con la misma información de las casillas recibidas el día de la jornada electoral, realizarán una estimación para la votación de cada uno de los candidatos a la Gubernatura del Estado de Nuevo León, así como, para el porcentaje de participación. Por lo anterior, para integrar los resultados anteriores se propone usar la unión de al menos dos de las tres estimaciones obtenidas. Sin embargo, la manera definitiva de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros.

5.2.1.1.4. Sinaloa³¹

Introducción

En un sistema de partidos altamente competitivo como el mexicano existen mecanismos que contribuyen a la certidumbre en los procesos electorales locales y federales. En este sentido los conteos rápidos son un mecanismo que da principalmente certeza, así como objetividad a los actores políticos en un proceso electoral.

Para el desarrollo de un conteo rápido es necesario que las autoridades electorales locales y federales, las Direcciones Ejecutivas de Organización Electoral, y de Capacitación Electoral y de Educación Cívica, del Registro Federal de Electores, las Juntas locales y distritales, los capacitadores asistentes electorales, supervisores electorales, funcionarios de casilla, consejos locales y distritales, los organismos públicos locales (OPL), así como los partidos políticos confluyan en el objetivo de realizar el Conteo Rápido para dar certeza al proceso electoral.

El Conteo Rápido es un procedimiento estadístico que permite la estimación de la tendencia de los resultados de la elección el mismo día de la jornada electoral, con cierta precisión y nivel de confianza a partir de una muestra aleatoria de las casillas de escrutinio y cómputo.

Para el diseño muestral del estado de Sinaloa se realizaron diferentes análisis y consideraciones técnicas que se discutieron en el Grupo 1, lo cual permitió definir el diseño muestral que se utilizará en el Conteo Rápido de la gubernatura del estado de Sinaloa el 6 de junio de 2021, que a continuación se presentan.

Estrategia de selección

Muestreo estratificado

El Conteo Rápido del proceso electoral 2020-2021 para la elección de la gubernatura del estado de Sinaloa se realizó con base a un procedimiento estadístico a partir de una muestra aleatoria del total de las casillas instaladas el

³¹ Elaborado por la Mtra. Claudia Esther Ortiz Guerrero con la colaboración del Dr. Vicente González Juárez

día de la jornada electoral, las cuales son aprobadas por los consejos Distritales del Instituto Nacional Electoral (INE). En el Conteo Rápido se estimarán las tendencias de los resultados finales de la jornada electoral. Un diseño muestral adecuado puede proporcionar rápidamente información con suficiente precisión y evaluar el margen de certidumbre de los resultados.

Bajo este contexto, un diseño muestral busca ubicar las casillas de la zona electoral donde se implemente y que todas tengan una probabilidad de ser seleccionadas en un proceso aleatorio mayor a 0.

Para el caso del Conteo Rápido de Sinaloa, el diseño consiste en un Muestreo Aleatorio Estratificado porque éste otorga estimaciones más precisas cuando la población se encuentra constituida por grupos que presentan características semejantes entre sí, pero diferentes de las de otros grupos, de tal manera que se reduce la variabilidad de las unidades.

La estratificación hace posible muestrear diferentes partes de una población en diferentes formas, lo cual puede constituir una mejora para emitir conclusiones sobre la tendencia de los resultados electorales.

Atendiendo a las consideraciones anteriores, el resumen de las características de método de muestro son las siguientes:

- **Tipo de Muestreo:** Muestreo Aleatorio Estratificado
- **Unidad Muestral:** Casillas
- **Tipo de Estimador:** Estimador de Razón

$$\hat{p}_i = \text{Votos_Partido} / \text{Votos_Total}$$

- **Nivel de confianza:** 95%
- **Tamaño de la Muestra:** Definida atendiendo la precisión, no respuesta y hora de llegada.
- **Asignación de la muestra:** Selección de muestra aleatoria proporcional al estrato:

$$n_h = \frac{N_h}{N} n$$

Donde:

n_h Es el número de casillas escrutadas en el h-ésimo estrado.

N_h Es el número total de casillas en el h-ésimo estrado

n Es el número total de la muestra

N Total de casillas en Sinaloa aplicando simulación.

Definición de estratos

Con base a los artículos 371, 372 y 373 del Reglamento de Elecciones (RE) se establece que el diseño, implementación y operación de los Conteos Rápidos deberán seguir criterios científicos, la teoría y los métodos de inferencia para realizar las estimaciones de los resultados de las elecciones, así como definir el diseño de la muestra. El marco muestral fue construido a partir del total de las casillas que se determine instalar para la jornada electoral.

Para la construcción de la muestra del Conteo Rápido del estado de Sinaloa para el proceso electoral 2020-2021 se consideró la experiencia y las recomendaciones contenidas en los *Informes Finales* de los Comités de Conteo Rápido (COTECORA) de años anteriores. En este sentido la estrategia se basa en la selección de las casillas de escrutinio y cómputo a partir de un muestreo aleatorio estratificado que abarca la mayor dispersión geográfica electoral posible como marca el RE.

En el Muestreo Aleatorio Estratificado, las unidades de la población se agrupan sobre la base de la similitud de alguna característica. Cada estrato se define a partir de algún factor que puede influenciar en el resultado del muestreo y que no puede ser sustraído. Los estratos deben ser exhaustivos (que abarquen a todas las unidades muestrales de tal manera que no quede ninguna sin formar parte de un estrato) y excluyentes (que no exista ninguna unidad muestral que forme parte de dos o más estrato a la vez).

Específicamente el artículo 373 del RE establece que la muestra debe abarcar la mayor dispersión geográfica posible y que éste es un factor de variabilidad estadística, los estratos considerados atienden a criterios geográficos. Por lo cual se consideraron los siguientes criterios de estratificación en el diseño muestral:

- Distritos Federales (7 Estratos)

- Distritos Locales (24 Estratos)
- Distritos Federales y Distritos Locales (35 Estratos)
- Distritos Federales y Distritos Locales colapsados (19 Estratos)
- Municipios (18 Estratos)

La selección del estrato más adecuado para el Conteo Rápido de la elección de la gubernatura en Sinaloa se realizó atendiendo a dos criterios: La precisión, que guarda estrecha relación en el error de estimación muestral el cual se pretende que sea alrededor del 1% y el porcentaje de capacitadores asistentes electorales (CAE) que participarán reportando 1 sola casilla para el Conteo Rápido.

En la Tabla 1, se muestra el análisis de las precisiones considerando diferentes tamaños de muestra.

Tabla 1 Análisis de precisiones para el Conteo Rápido a la Gubernatura de Sinaloa

n	Federal	Local	Federal-Local	19 Estratos
200	1.52	1.44	1.31	1.42
250	1.33	1.28	1.23	1.24
300	1.23	1.12	1.10	1.11
350	1.19	1.02	1.03	1.04
400	1.05	0.94	0.98	1.02
450	0.97	0.89	0.93	0.88
500	0.94	0.84	0.86	0.86
550	0.86	0.84	0.80	0.81
600	0.84	0.75	0.74	0.77
650	0.79	0.74	0.72	0.76
700	0.79	0.70	0.68	0.72
750	0.74	0.67	0.68	0.68
800	0.73	0.66	0.64	0.63
850	0.67	0.62	0.63	0.61
900	0.67	0.62	0.60	0.59

Fuente: Elaboración propia con base a la información estadística sobre Elección gubernatura 2015, casillas aprobadas al 30 de marzo de 2021 (INE 2021).

Aunque el criterio más intuitivo para estratificar geográficamente podría ser por municipios, este criterio se ha dejado de lado porque para el caso de Sinaloa no es un criterio adecuado para reducir la variabilidad en este ejercicio debido a

que se aprobaron por la legislatura local dos municipios más en la entidad en marzo de 2021, y podría causar controversias en el plano social y político. De esta manera, se emplea la geografía electoral local y federal con la que dispone el INE la cual genera estratos más homogéneos.

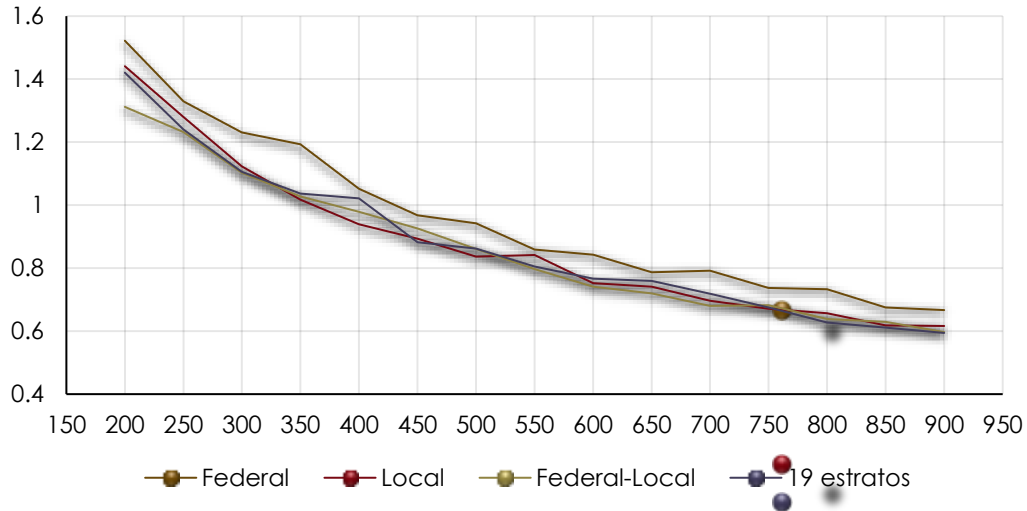
Tabla 2 Análisis de Precisiones para el Estrato Municipal

n	Municipal
350	1.032
400	0.933
450	0.860
500	0.837
550	0.802
600	0.742

Fuente: Alonso y Coria, Alberto y Morales Bolio Joaquín, con base a la información estadística sobre Elección gubernatura 2015, casillas aprobadas al 30 de marzo de 2021 (INE 2021).

De acuerdo con el análisis de precisiones, se puede apreciar que la precisión deseada se obtiene para una n de 400 cuando se estratifica bajo los criterios de Distrito Local y la combinación de Distrito Local y Distrito Federal. El resto de los estratos cuentan con un comportamiento similar siendo la estratificación que considera 19 estratos colapsados de la combinación entre Distrito Local y Distrito Federal la presenta una mejor disminución conforme el tamaño de la muestra aumenta. Se visualiza en la Gráfica 1.

Gráfica 1 Análisis de Precisión para las estratificaciones



Por otro lado, Distrito Federal y Distrito Local genera tamaños de muestra muy pequeños además considerando el principio de parsimonia se busca representar la información el menor número de estratos afectados, dada que el incluir una amplia cantidad de estratos puede traer problemas en la proporción de muestra asignada en cada uno lo que provocaría una baja representatividad en un grupo que se traduce en sesgos a en las estimaciones.

Participación de los capacitadores asistentes electorales (CAE)

Por otro lado, dada la inquietud en las múltiples funciones que desarrollan los CAE, en especial el día de la jornada electoral y en los últimos ejercicios de Conteo Rápido, se analizó también la proporción de CAE a los que les tocaría reportar 1, 2 y 3 o más casillas. Para este análisis se consideró la información de la elección a la gubernatura en 2015. Los resultados se muestran en la Tabla 3.

Tabla 3 Porcentajes de CAE con 1, 2, y 3 o más casillas

n	Federal			Local			Federal-Local Colapsado		
	1	2	3 o más	1	2	3 o más	1	2	3 o más
200	93.1	6.1	0.8	92.0	6.8	1.3	92.6	6.2	1.2
250	91.6	7.7	0.8	90.5	8.2	1.4	91.0	7.8	1.3
300	90.1	9.1	0.8	88.9	9.6	1.5	89.6	9.0	1.4
350	88.5	10.5	1.0	87.2	11.1	1.7	88.2	10.3	1.5

n	Federal			Local			Federal-Local Colapsado		
	1	2	3 o más	1	2	3 o más	1	2	3 o más
400	87.0	11.9	1.1	85.8	12.4	1.8	86.6	11.8	1.6
450	85.4	13.3	1.3	84.2	13.7	2.1	85.2	13.0	1.8
500	83.8	14.6	1.1	82.6	15.0	2.4	83.8	14.2	2.0
550	82.3	15.8	1.8	81.1	16.3	2.6	82.2	15.5	2.3
600	80.7	17.2	2.1	79.5	17.6	3.0	80.8	16.7	2.5
650	79.2	18.4	2.4	77.9	18.8	3.3	79.2	18.0	2.8
700	77.6	19.7	2.8	76.4	19.9	3.7	77.7	19.1	3.2
750	76.1	20.8	3.2	74.7	21.1	4.2	76.4	20.1	3.5
800	74.5	22.0	3.6	73.2	22.2	4.6	74.8	21.4	3.8

Fuente: Elaboración propia con base a la información estadística sobre Elección gubernatura 2015, casillas aprobadas al 30 de marzo de 2021 (INE 2021).

De acuerdo con la Tabla 3, el porcentaje de CAE a los cuales les corresponde reportar 1 sola casilla es bastante alto para todos los niveles de muestra reportados, ésta está por encima del 85% para muestras de 450 y superior al 75% para muestras de 750. Cuando se emplea la estratificación Distrito Local y Distrito Federal Colapsado el porcentaje es ligeramente mejor que en los otros criterios.

Partiendo de los criterios de precisión y de los porcentajes de los CAE con 1 casilla se ha determinado como criterio de estratificación Distrito Local y Distrito Federal Colapsado por lo que el diseño de muestreo para el Conteo Rápido en Sinaloa consta de 19 estratos.

Otro variable que influye en la determinación del tamaño de muestra de 750 casillas se basa en la información del proceso de llegada de los datos y no respuesta del Conteo Rápido de 2018, que a continuación se expone.

El proceso de llegada de los datos y no respuesta

Otro criterio importante para determinar el tamaño de la muestra en los ejercicios de Conteo Rápido es la prontitud en la cual se pueden emitir resultados preliminares. Estos estarán oportunamente disponibles en la medida de que la información recolectada en las casillas por los CAE puede ser transmitida de manera eficiente a los centros de captura de datos en las Juntas Distritales.

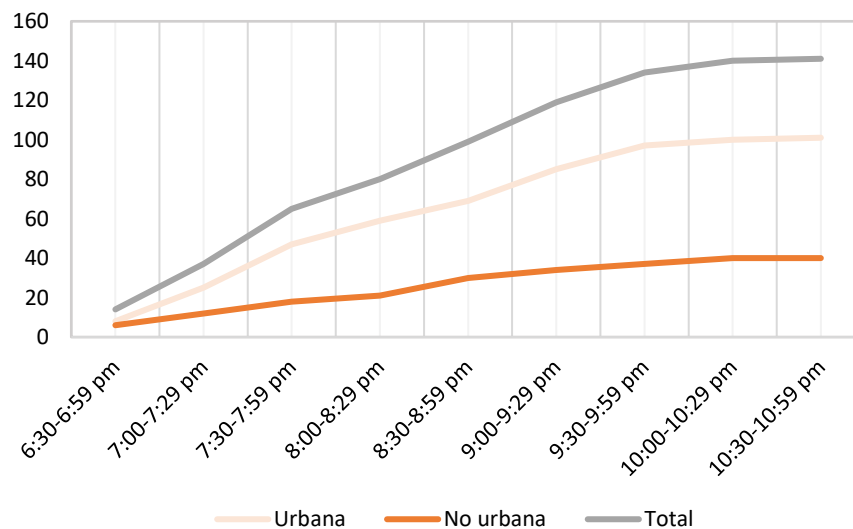
Por ello se toma como base la recepción de la muestra por hora de la elección presidencial de 2018 asociados al estado de Sinaloa para determinar el porcentaje de no respuesta el cual se puede resumir en la tabla 4.

Tabla 4. Casillas totales reportadas y no reportadas de la muestra del COTECORA de la elección presidencial 2018 Estado de Sinaloa

Casillas	Urbanas	No Urbanas	Total
Reportadas	101	40	141
No Reportadas	60	30	90
Total	161	70	231

Como se puede apreciar en la Tabla 4. Existe una alta cantidad de No respuesta, de la muestra estimada de 231 al cierre, 11 de la noche, sólo se habían recibido 141 correspondiente al 60.2%. De las casillas urbanas sólo se reportó el 62%, mientras que de las casillas no urbanas se reportó el 57%. Esta tendencia se visualiza en la Gráfica 2.

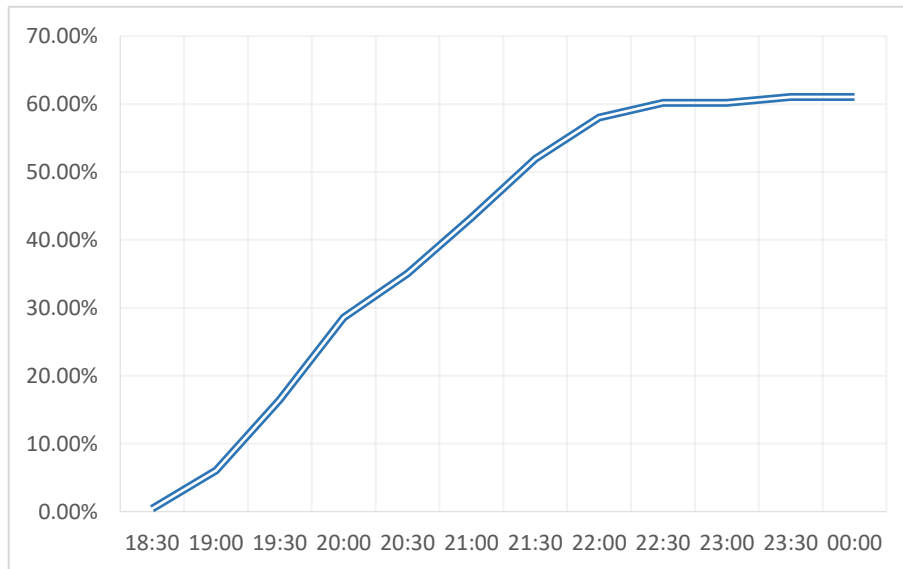
Gráfica 2 Casillas Reportadas Acumuladas



En la Gráfica 2, se puede apreciar cómo fue la entrega de los reportes de las casillas electorales conforme avanza la jornada de acuerdo con el tipo de casilla (Urbana o no). Notando una baja entrega en ambos rubros. En porcentaje máximo de entrega de 60.2% se alcanza a las 10:30 de la noche, pero las horas previas ya se podía apreciar una disminución en la recepción de los reportes de

las casillas. Por su parte, el porcentaje de no entrega acumulado se encuentra plasmado en la gráfica 3.

Gráfica 3 Porcentaje acumulado de recepción de la muestra por hora



Determinación del tamaño de muestra

El día de la jornada electoral estaríamos esperando recibir alrededor de las 11:00 pm el 60% de la muestra solicitada. Por lo tanto, la muestra requerida para obtener 400 casillas efectivas es de $n = 750$, al cual se le asocia un error de 0.68% y un porcentaje de CAE con una sola casilla del 76%. Por lo cual, este tamaño permitirá tener estimaciones precisas, en un tiempo oportuno, con tolerancia a la no respuesta y sin imponer una jornada excesiva a los CAE.

No obstante, procedente del análisis de no respuesta del Contero Rápido de las elecciones anteriores, se determinó que, con el propósito de garantizar dichos criterios se mantengan, es necesario aumentar el tamaño de la muestra.

Asignación de la muestra

Las 750 casillas consideradas para el estado de Sinaloa se agruparon en los 19 estratos con una asignación proporcional al tamaño de las casillas. La

información sobre el total de casillas corresponde a la Dirección de Ejecutiva de Organización Electoral del 30 de marzo de 2021, se dispone de un total de **4,982** casillas aprobadas por los Consejos Distritales. La distribución de las casillas por estrato, así como la asignación de la muestra en cada uno de estos se resume en la Tabla 5.

Tabla 5 Muestra Sinaloa Colapsos de distritos locales con intercepción con distritos federales (Estratos)

Estrato	Dist. Federal	Dist. Locales	Nh	nh
1	1	20 y 23	285	43
2	1	21 y 22	167	26
3	1	24	185	28
4	2	2	183	28
5	2	3	187	28
6	2	4 y 5	321	48
7	3	6	189	28
8	3	9	281	42
9	3	10	238	36
10	3	11	176	26
11	4	1	262	39
12	4	4 y 8	259	39
13	4	6 y 7	253	38
14	5	12 15 y 16	409	62
15	5	13 y 14	211	32
16	6	18 y 19	344	52
17	6	20 21 22 y 23	346	52
18	7	12 13 y 14	333	50
19	7	15 17 y 18	353	53
		Totales:	4,982	750

Fuente: Elaboración propia con base a la información estadística sobre Elección gubernatura 2015, casillas aprobadas al 30 de marzo de 2021 (INE 2021).

Algunas consideraciones

Se debe hacer énfasis en que este es un diseño muestral basado en la información de las bases de datos de la elección de gobernador de 2015 y de la elección presidencial de 2018 (INE 2021), por lo cual las estimaciones y precisiones tienen la finalidad de determinar el tamaño de muestra adecuado desde la

perspectiva teórica para obtener el nivel de precisión y de confianza estimado. Sin embargo, como la realidad social y política es cambiante, las variables pueden modificarse, como son la tasa de no respuesta que esperamos sea mayor a la presentada en el Conteo Rápido de 2018. Al término de la elección se estimará la precisión obtenida.

Los resultados de las simulaciones presentadas son válidos únicamente para el método de estimación presentado bajo los criterios y condiciones descritas. Se pueden emplear métodos de estimación alternativos que arrojarán precisiones distintas o bien que generan un tamaño de muestra estimada diferente.

Existen fórmulas analíticas para calcular la varianza y por lo tanto el intervalo de confianza para el estimador de razón combinado. Sin embargo, siempre se utiliza la aproximación normal para construir el intervalo de confianza y esto puede ocasionar problemas cuando se estiman porcentajes muy pequeños. En estos casos, el intervalo inferior suele ser negativo y se debe truncar el intervalo en el cero. Esta solución, no tiene ningún fundamento técnico. El *Bootstrap* soluciona este tipo de problemas de manera automática.

La metodología aquí expuesta está basada en consideraciones estadísticas teóricas para condiciones ideales, pero el día de la elección en circunstancias reales, pueden existir factores no contemplados que influyan directamente en los resultados estimados.

Hay que mencionar, por un lado, que, si se recibe un tamaño de muestra menor a 750 casillas, el día de la Jornada Electoral, todavía se pueden tener estimaciones confiables si la distribución de la muestra recibida es adecuada. Sin embargo, las precisiones serán menores como se puede ver en la Tabla 1. Por ejemplo, aún para un tamaño de muestra de 300 casillas con una composición y distribución suficiente desde el punto de vista teórico y que corresponda al diseño muestral aquí definido, se alcanzaría una precisión alrededor de 1%, pero se podría tener un sesgo importante en las estimaciones reportadas. Por otro lado, también se debe señalar que si se recibe una muestra menor a 300 casillas con una composición y calidad muy alejada de la definimos mediante el diseño muestral será difícil obtener resultados que den certidumbre en las estimaciones.

Finalmente hay que señalar que los diferentes procedimientos de estimación que implementarán los tres integrantes del Grupo 1 del COTECORA para la realización del estado de Sinaloa estarán basados en el Diseño Muestral definido en esta sección.

Integración de las estimaciones

Cada uno de los tres integrantes del Grupo 1 del COTECORA, con la misma información de las casillas recibidas el día de la jornada electoral, realizará una estimación para la votación de cada una de las candidaturas a la gubernatura del estado de Sinaloa, así como, para el porcentaje de participación ciudadana. Por lo anterior, para integrar los resultados anteriores se propone usar la unión de al menos dos de las tres estimaciones obtenidas. Sin embargo, la manera definitiva de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros.

5.2.1.1.5. Tlaxcala ³²

I.Elementos para diseño muestral

a. El parámetro de precisión

Los conteos rápidos constituyen un instrumento para que la ciudadanía y actores interesados puedan conocer en forma preliminar un primer estimado del resultado de la jornada electoral. Los resultados se expresan no como estimaciones puntuales sino como intervalos en el porcentaje de votación cada uno con una confianza determinada. Ciertamente no se sabe cuál será el sentido de voto de la población en la próxima Jornada Electoral y por ello, un criterio de relevancia en el diseño es el tratar de establecer un tamaño de muestra tal que se espere que los intervalos de estimación sean reducidos y que por tanto disminuya la probabilidad de que exista un traslape en la noche de la Jornada Electoral. Para hacer estos estudios se analizaron distintos esquemas de estratificación del estado y de tamaños de la muestra.

1. Escenarios de estratificación explorados

Para la realización de estos trabajos se utilizaron los resultados electorales en la última elección para gobernador del estado y de los resultados, restringidos a esta entidad federativa, en la elección de Presidente de la República 2018. Los parámetros básicos de estas elecciones se presentan a continuación.

Elección	Distritos		Casillas				
	Federales	Locales	Urbanas	Mixtas	Rurales	Mixtas + Rurales	Total casillas
Gobernador 2016	3	15	1,119	258	124	382	1,501
Presidente 2018			1,130	275	128	403	1,533

Tabla 1*

* Elaborado con base en los resultados electorales proporcionados por el INE

Primeramente, se exploraron las siguientes estratificaciones³³:

- Distritos federales.

³² Elaborado por el Dr. Alberto Alonso y Coria con la colaboración de Fis. Joaquín Morales Bolio

³³ Se utilizó el marco cartográfico vigente en 2020

- Distritos Locales.
- Distritos Locales y Federales separando la parte rural y urbana. (explorándose como criterio de separación el 30% y 40%. Esto es, bifurcando el estrato sólo si el distrito contaba con al menos ese porcentaje de casillas rurales o no-urbanas).

Para cada uno de los escenarios y cada uno de los tamaños de muestra se realizaron mil muestras distintas y se obtuvo el promedio de la precisión. El promedio que se reporta a continuación es la peor precisión (la más grande) para cada uno de los partidos o en su caso coaliciones que compitieron en la elección en cuestión. Los cortes del 30% o 40% en sólo se usaron de manera indicativa para la posible exploración de un corte más fino. No todos los distritos locales o federales se separaban en subestratos urbano-mixto-rural (o urbano y rural-mixto) en razón de la inviabilidad práctica de la división por no haber suficientes casillas de alguno de los tipos que diera lugar a estratos suficientemente poblados en términos de número de casillas. Se anexa una tabla que sirvió de guía para determinar en qué distritos si se consideró viable la exploración rural-urbano (la tabla presentada es la correspondiente a la elección de gobernador. Una semejante se utilizó en el caso del análisis de la elección de presidente).

Distrito / Estrato	Tlaxcala											
	Distritos Federales			Distritos Locales			Manual 1			Manual 2		
	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas	Casillas x estrato	% de casillas Rurales	% de casillas Rurales + Mixtas
1	502	15.14%	39.24%	94	7.45%	61.70%	203	16.26%	56.65%	94	7.45%	61.70%
2	484	1.24%	5.58%	108	27.78%	57.41%	320	13.13%	34.38%	108	27.78%	57.41%
3	501	8.58%	32.93%	91	13.19%	45.05%	209	4.78%	22.01%	112	10.71%	40.18%
4				100		3.00%	201		1.99%	100		3.00%
5				97	26.80%	57.73%	182	18.68%	47.80%	109	23.85%	52.29%
6				111	8.11%	25.23%	205	1.46%	10.24%	111	8.11%	25.23%
7				112			167	1.80%	3.59%	118		
8				96	3.13%	8.33%				167	1.80%	3.59%
9				104		2.88%				89	7.87%	24.72%
10				89	7.87%	24.72%				93	29.03%	69.89%
11				93	29.03%	69.89%				113	2.65%	5.31%
12				103	1.94%	1.94%				83		4.82%
13				99	1.01%	8.08%				98	1.02%	18.37%
14				98	1.02%	18.37%				92		16.30%
15				92		16.30%						

Tabla 2*

* Elaborado con base en las casillas para la elección de gobernador con el marco cartográfico vigente. Marco proporcionado por el INE

Un punto de relevancia en el diseño fue la consideración de que una submuestra de la que se plantea utilizar para Tlaxcala pudiera ser utilizada para el ejercicio de conteo rápido para la estimación de la conformación de la cámara de diputados. Así pues, se consideró pertinente que los estratos estuvieran contenidos en los distritos electorales federales en razón de que estos serán utilizados para ese ejercicio.

En una segunda etapa se exploraron distintas agrupaciones de distritos locales. Las dos más notables se describen a continuación:

TLAXCALA							
AGRUPACIÓN 1				AGRUPACIÓN 2			
Estrato	Intersección			Estrato	Intersección		
	Distrito Electoral Federal		Distrito Electoral local		Distrito Electoral Federal	Distrito Electoral local	
1	3	y	5, 1	1	3	Y	1
	3	y	8	2	1	Y	2
2	1	y	2, 3	3	1	Y	3
	1	y	4	4	1	Y	8
	1	y	8		5	3	Y
3	3	y	6, 14	6	3	Y	8
	3	y	13		7	2	Y
	3	y	12	8	2	Y	8
4	2	y	13, 7	9	1	Y	10
5	1	y	11, 10	10	1	Y	11
6	2	y	12, 15	11	2	Y	12
7	2	y	8, 9	12	3	Y	12
				13	3	Y	13
				14	3	Y	14
				14	2	Y	15

Tabla 3

Como puede observarse la llamada agrupación 2 consiste fundamentalmente en los distritos locales, adecuando al estrato para resolver los traslapes con los distritos federales.

2. Precisión vs. Tamaño de muestra.

La siguiente tabla presenta el análisis correspondiente a la elección de gobernador 2016.

Tlaxcala. Gobernador 2016

No. Estratos	3	5	15	20	18	7	10	9	14	19	17	
Tamaño de la muestra	Distritos Federales		Distritos Locales				Agrupación 1			Agrupación 2		
	Distritos Federales	con substratos mixtos 30%	Distritos Locales	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 1	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 2	con substratos mixtos 30%	con substratos mixtos 40%	
130	1.39	1.40	1.25	1.21	1.22	1.34	1.33	1.35	1.25	1.24	1.24	
140	1.34	1.31	1.19	1.18	1.19	1.30	1.29	1.29	1.20	1.20	1.20	
150	1.29	1.28	1.16	1.15	1.16	1.25	1.25	1.25	1.18	1.14	1.15	
160	1.25	1.25	1.11	1.09	1.09	1.20	1.20	1.19	1.09	1.10	1.11	
170	1.20	1.19	1.07	1.06	1.05	1.17	1.17	1.16	1.09	1.08	1.06	
180	1.16	1.16	1.04	1.02	1.03	1.15	1.13	1.13	1.05	1.04	1.04	
190	1.13	1.13	1.00	1.01	1.00	1.09	1.09	1.09	1.02	1.00	1.01	
200	1.10	1.10	0.98	0.96	0.98	1.07	1.06	1.05	1.01	0.99	0.98	
210	1.06	1.06	0.94	0.93	0.95	1.03	1.02	1.04	0.96	0.94	0.95	
220	1.03	1.03	0.93	0.92	0.92	1.01	1.01	1.00	0.93	0.93	0.91	
230	1.01	1.01	0.91	0.89	0.90	0.98	0.98	0.99	0.92	0.91	0.91	
240	0.98	0.99	0.89	0.88	0.87	0.95	0.96	0.95	0.89	0.88	0.87	
250	0.96	0.96	0.87	0.86	0.86	0.93	0.93	0.93	0.87	0.86	0.86	
260	0.93	0.93	0.85	0.83	0.83	0.91	0.91	0.91	0.85	0.84	0.84	
270	0.92	0.92	0.82	0.81	0.81	0.90	0.88	0.89	0.83	0.82	0.82	
280	0.90	0.89	0.80	0.80	0.79	0.88	0.87	0.87	0.81	0.80	0.81	
290	0.88	0.88	0.79	0.77	0.79	0.85	0.85	0.85	0.80	0.79	0.79	
300	0.86	0.85	0.77	0.76	0.77	0.84	0.83	0.83	0.77	0.76	0.77	

Tabla 4. precisión elección gobernador 2016

Como puede observarse la estratificación con la agrupación 2, que es fundamentalmente a través de distritos locales, pero teniendo en cuenta su intersección con los distritos federales, con una muestra de 210 casillas genera una precisión inferior al uno por ciento.

En un ejercicio semejante realizado con los datos de la elección presidencial 2018, si bien coincidentemente apuntan a la bondad de la estratificación vía el escenario 2, es a partir de un tamaño de muestra de 140 casillas que se obtendría una precisión inferior al uno por ciento. (Tabla 5).

TLAXCALA. Datos elección federal

No. Estratos	3	5	15	20	17	20	19	7	8	10	9	14	18	16	19	18
Tamaño de la muestra	Distritos Federales		Distritos Locales				Agrupación 1				Agrupación 2					
	Distritos Federales	con substratos mixtos 30%	Distritos Locales	con substratos rurales 30%	con substratos rurales 40%	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 1	con substratos rurales 30%	con substratos mixtos 30%	con substratos mixtos 40%	Agrupación 2	con substratos rurales 30%	con substratos rurales 40%	con substratos mixtos 30%	con substratos mixtos 40%
130	1.16	1.14	1.02	1.02	1.01	1.01	1.02	1.09	1.08	1.07	1.07	1.01	1.01	1.04	1.01	1.02
140	1.14	1.09	0.97	0.96	0.98	0.98	0.97	1.05	1.05	1.04	1.04	0.99	0.99	1.00	0.98	0.99
150	1.10	1.05	0.96	0.94	0.94	0.95	0.94	1.01	1.01	1.00	0.99	0.97	0.96	0.99	0.96	0.96
160	1.06	1.01	0.91	0.91	0.90	0.90	0.90	0.96	0.97	0.98	0.97	0.92	0.92	0.91	0.91	0.91
170	1.03	0.99	0.89	0.88	0.89	0.89	0.88	0.94	0.94	0.94	0.93	0.89	0.89	0.88	0.89	0.88
180	0.99	0.95	0.85	0.84	0.84	0.83	0.85	0.91	0.91	0.89	0.90	0.87	0.87	0.86	0.84	0.85
190	0.97	0.93	0.81	0.82	0.81	0.81	0.83	0.89	0.89	0.88	0.88	0.83	0.84	0.83	0.84	0.82
200	0.93	0.89	0.80	0.80	0.81	0.81	0.81	0.85	0.86	0.84	0.86	0.81	0.80	0.82	0.81	0.81
210	0.91	0.87	0.78	0.77	0.79	0.78	0.78	0.83	0.82	0.82	0.82	0.78	0.77	0.79	0.77	0.77
220	0.89	0.86	0.76	0.75	0.76	0.76	0.75	0.81	0.81	0.80	0.81	0.77	0.76	0.77	0.77	0.77
230	0.87	0.83	0.75	0.73	0.75	0.73	0.73	0.80	0.79	0.79	0.79	0.75	0.74	0.74	0.73	0.74
240	0.84	0.81	0.73	0.71	0.72	0.72	0.72	0.77	0.76	0.77	0.77	0.73	0.73	0.73	0.72	0.73
250	0.83	0.79	0.71	0.71	0.70	0.70	0.70	0.75	0.74	0.75	0.74	0.72	0.71	0.70	0.71	0.71
260	0.80	0.77	0.70	0.69	0.68	0.69	0.69	0.73	0.74	0.73	0.73	0.70	0.69	0.69	0.69	0.70
270	0.79	0.75	0.67	0.67	0.67	0.67	0.67	0.72	0.72	0.71	0.72	0.68	0.67	0.68	0.68	0.68
280	0.76	0.74	0.65	0.66	0.66	0.65	0.65	0.70	0.70	0.70	0.70	0.66	0.66	0.66	0.65	0.67
290	0.75	0.72	0.64	0.64	0.64	0.65	0.64	0.68	0.69	0.69	0.69	0.65	0.65	0.65	0.65	0.65
300	0.73	0.71	0.63	0.63	0.63	0.62	0.63	0.67	0.67	0.67	0.67	0.64	0.64	0.63	0.63	0.64

Tabla 5. Precisión elección federal 2018

3. Número de casillas por estrato

Un parámetro a considerar es el número de casillas muestra por estrato. Éste es relevante, en tanto que es posible que por vicisitudes del día de la jornada electoral pudiera no ser posible la obtención de la información de alguna de ellas. Es por ello que es de importancia tener en cuenta la siguiente tabla:

Tamaño de la muestra	Tlaxcala											
	Distritos Federales			Distritos Locales			Agrupación 1			Agrupación 2		
	Distritos Federales	con subestratos mixtos 30%	con subestratos mixtos 40%	Distritos Locales	con subestratos mixtos 30%	con subestratos mixtos 40%	Agrupación 1	con subestratos mixtos 30%	con subestratos mixtos 40%	Agrupación 2	con subestratos mixtos 30%	con subestratos mixtos 40%
130	42 - 44	14 - 42	42 - 44	8 - 10	2 - 10	4 - 10	15 - 28	8 - 18	8 - 28	7 - 15	2 - 15	4 - 15
140	46 - 47	16 - 46	46 - 47	8 - 11	3 - 11	4 - 11	16 - 30	8 - 20	8 - 30	8 - 16	3 - 16	4 - 16
150	49 - 51	17 - 49	49 - 51	9 - 11	3 - 11	4 - 11	17 - 32	9 - 21	9 - 32	8 - 17	3 - 17	5 - 17
160	52 - 54	18 - 52	52 - 54	10 - 12	3 - 12	4 - 12	18 - 34	9 - 23	9 - 34	9 - 18	3 - 18	5 - 18
170	55 - 57	19 - 55	55 - 57	10 - 13	3 - 13	5 - 13	19 - 37	10 - 24	10 - 37	9 - 19	3 - 19	5 - 19
180	59 - 61	20 - 59	59 - 61	11 - 14	3 - 14	5 - 14	20 - 39	11 - 25	11 - 39	10 - 20	3 - 20	5 - 20
190	62 - 64	21 - 62	62 - 64	11 - 14	4 - 14	5 - 14	21 - 41	11 - 27	11 - 41	11 - 21	4 - 21	6 - 21
200	65 - 68	22 - 65	65 - 68	12 - 15	4 - 15	6 - 15	22 - 43	12 - 28	12 - 43	11 - 22	4 - 22	6 - 22
210	68 - 71	23 - 68	68 - 71	13 - 16	4 - 16	6 - 16	24 - 45	12 - 30	12 - 45	12 - 24	4 - 24	6 - 24
220	72 - 74	24 - 72	72 - 74	13 - 17	4 - 17	6 - 17	25 - 47	13 - 31	13 - 47	12 - 25	4 - 25	7 - 25
230	75 - 78	26 - 75	75 - 78	14 - 17	4 - 17	6 - 17	26 - 49	13 - 32	13 - 49	13 - 26	4 - 26	7 - 26
240	78 - 81	27 - 78	78 - 81	14 - 18	5 - 18	7 - 18	27 - 52	14 - 34	14 - 52	13 - 27	5 - 27	7 - 27
250	81 - 84	28 - 81	81 - 84	15 - 19	5 - 19	7 - 19	28 - 54	15 - 35	15 - 54	14 - 28	5 - 28	8 - 28
260	85 - 88	29 - 85	85 - 88	16 - 20	5 - 20	7 - 20	29 - 56	15 - 37	15 - 56	15 - 29	5 - 29	8 - 29
270	88 - 91	30 - 88	88 - 91	16 - 20	5 - 20	7 - 20	30 - 58	16 - 38	16 - 58	15 - 30	5 - 30	8 - 30
280	91 - 95	31 - 91	91 - 95	17 - 21	5 - 21	8 - 21	31 - 60	16 - 40	16 - 60	16 - 31	5 - 31	8 - 31
290	94 - 98	32 - 94	94 - 98	17 - 22	5 - 22	8 - 22	33 - 62	17 - 41	17 - 62	16 - 33	5 - 33	9 - 33
300	98 - 101	33 - 98	98 - 101	18 - 23	6 - 23	8 - 23	34 - 65	18 - 42	18 - 65	17 - 34	6 - 34	9 - 34

Tabla 6.*

* Datos de la elección gobernador 2016

b. La situación con Capacitadores Asistentes Electorales (CAE)

Una limitante para el ejercicio del conteo rápido es la necesidad de un intenso trabajo en campo durante la noche del día de celebración de los comicios. En razón de que la muestra de casillas se genera en cada estrato en forma aleatoria, es posible que algunos CAE tengan la responsabilidad de reportar más de una casilla. Así pues, un parámetro de relevancia consiste en el estimar el porcentaje de CAE que tendrán sólo una casilla para reportar y que por tanto puedan acudir en forma oportuna a su ubicación y proceder, conforme a los lineamientos, a informar los resultados. Quizá sea importante mencionar que en estudios realizados por organización electoral³⁴ no existe una relación totalmente directa en la oportunidad de reporte y el número de casillas a atender por el CAE. Sin embargo, es claro que se facilita la operación en campo si se reduce el número

³⁴ 10feb2021_DEOE_PRINCIPALES RESULTADOS CR 2018 y 2019.pptx. De hecho, sólo el 70% de los CAE con una sola casilla reportaron a tiempo para el conteo rápido federal.

de casillas a atender por esta figura y por ende su estudio formó parte de los parámetros a considerar.

Tlaxcala	
Tamaño de la muestra	Porcentaje de CAEs con una casilla de la muestra
130	87.59
140	87.19
150	85.60
160	84.62
170	84.01
180	82.70
190	81.48
200	80.64
210	79.51
220	78.58
230	77.56
240	76.95
250	75.96
260	74.90
270	73.42
280	72.58
290	71.51
300	70.11

Tabla 7*

*Datos elección gobernador

c. Arribo de la muestra en el caso de la elección presidencial 2018

A diferencia del PREP que por su naturaleza en los primeros minutos u horas tiene un posible sesgo geográfico producido por la distinta velocidad de llegada de las actas a las sedes distritales entre, digamos, las casillas urbanas y rurales, el conteo rápido cuenta con un medio más eficiente de comunicación que minimiza o disminuye esta problemática. En razón de la posible existencia de una comunicación menos eficiente en las casillas rurales, se estudió el posible efecto en la hora de llegada de la información de las casillas en la muestra. Se realizó un estudio sobre la forma de arribo de la información de las casillas muestra utilizando los datos correspondientes a Tlaxcala del conteo rápido celebrado para la elección presidencial. Los datos y gráficas asociadas se presentan a continuación.

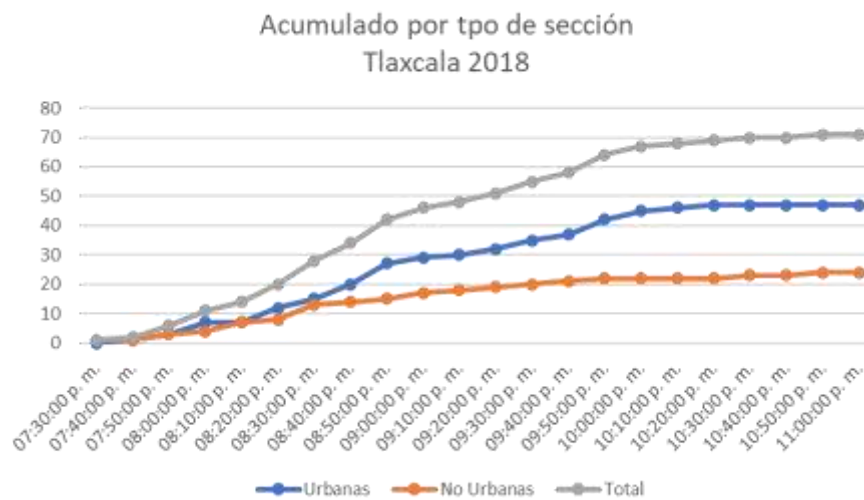
Datos Tlaxcala 2018				
	Urbanas	Mixtas	Rurales	Mixtas + Rurales
Total de casillas ¹	1130	275	128	403
	73.71%	17.94%	8.35%	26.29%
Casillas marco muestral ³	48	16	9	25
	65.75%	21.92%	12.33%	34.25%
Casillas capturadas ²	47	16	9	25
	65.28%	22.22%	12.50%	34.72%

Tabla 8

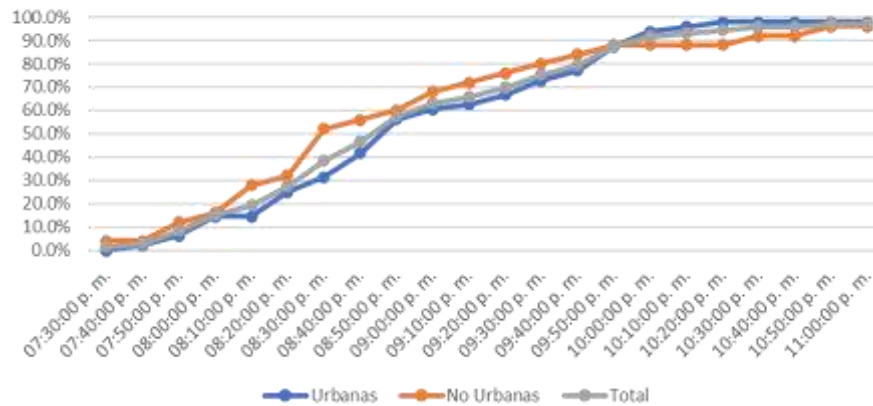
¹ Listado_casillas_2018.csv

² Corte TOTALES0100020040-txt 2 jul 2018 12:40 a.m.

³ 4-ConteoRapido18MUESTRA-ELECCION-PRESIDENCIAL.csv



Porcentaje de Llegada por tipo de sección
Tlaxcala 2018



Acorde a la gráfica inferior derecha, es de destacar que Tlaxcala es un estado que en el conteo rápido presidencial 2018 no presentó una diferencia real en la velocidad de captura entre casillas urbanas y no-urbanas. Esto ciertamente abona a no incluir en la estratificación posibles subestratos.

d. Estratificación recomendada

Del análisis de los datos obtenidos se concluye que la estratificación por la agrupación 2 es un buen esquema muestral amén de que todos los estratos están contenidos en algún distrito electoral federal. La tabla resumen siguiente nos indica que el tamaño de muestra de 180 casillas sería el adecuado para realizar la estimación.

TLAXCALA			
No. Estratos = 14			
Precisión (%)			
	Elección Gobernador		Elección Federal
Tamaño de la muestra	Criterio Razón	Criterio remuestreo local*	Criterio Razón
170	1.09	1.16	0.89
180	1.05	1.11	0.87
190	1.02	1.13	0.83
200	1.01	1.06	0.81
210	0.93	1.04	0.78

Tabla 9

* Datos proporcionados por Dr. Gabriel Nuñez y M. en C. Emiliano Geneyro

e. Expansión por arribo.

Un parámetro a considerar es el arribo de la muestra. En el ejercicio realizado en el proceso electoral 2017-18 en las entidades que tuvieron elección de gobernador se pudo observar que de la muestra solicitada arribó un porcentaje bajo. Como puede observarse en la tabla siguiente el mayor porcentaje de muestra recibida fue el de Morelos con 80.0% y el peor fue el de Tabasco con un 58.2%.³⁵

³⁵ Informe COTECORA junio 2018.

Conteos rápidos elecciones Gobernador 2018

Entidad	Porcentaje recibido	Casillas totales	Casillas en muestra	Muestra recibida
cdmx	58.4	12,976	1,108	647
chiapas	59.2	6,316	500	294
Guanajuato	71.4	7,463	500	357
jalisco	73.7	9,820	472	348
Morelos	80	2,421	200	160
Puebla	73.1	7,547	509	372
Tabasco	58.4	2,912	450	263
Veracruz	70.1	10,595	1,100	771
Yucatán	62.7	2,666	300	188
Total		62,716	5,139	3,400

Tabla 10*

* informe final COTECORA

Así pues, se considera en el diseño, que sólo arribará cerca del 67% de la muestra y por tanto se establece como parámetro de una muestra efectiva $180/.67 \sim 270^{36}$ casillas. Si bien la fracción muestral es alta, esto es derivado de que es la cuarta entidad más pequeña del país en término de número de casillas. El tamaño permite a su vez estimar que más del 73% de los CAE tendrán que reportar sólo una casilla.

f. Tabla final de Estratos

Para establecer la tabla final de estratos se utilizó el listado de casillas correspondiente al estado de Tlaxcala, aprobadas por los consejos distritales con corte al 30 de marzo de 2021. Se utilizará el escenario marcado como agrupación 2.

Estratificación: Agrupación de distritos locales atendiendo a distritos federales.

Precisión estimada: 0.87% a 1.11%

³⁶ Redondeado a la decena más próxima.

Tlaxcala						
ESTRATO	Entidad	Distrito Federal	Distrito Local	Nh	nh	
1	29	3	1	99	17	
2	29	1	2	113	19	
3	29	1	3	121	21	
			8			
4	29	1	4	103	18	
5	29	3	5	116	20	
			8			
6	29	3	6	119	20	
7	29	2	7	135	23	
			13			
8	29	2	8	176	30	
			9			
9	29	1	10	97	17	
10	29	1	11	99	17	
11	29	2	12	102	17	
12	29	3	12	94	16	
			13			
13	29	3	14	105	18	
14	29	2	15	100	17	
				TOTAL	1579	270

Tabla 11

II. Algunas consideraciones

1. Se debe señalar que la precisión en las estimaciones se maneja tanto en la etapa de planeación del ejercicio como en la presentación de los resultados finales del mismo. En la etapa de planeación se usa para determinar un tamaño de muestra adecuado teóricamente, para alcanzar el nivel de precisión y confianza deseadas en las estimaciones. Sin embargo, debido a que se fija este tamaño de muestra tomando como referencia los resultados de elecciones previas y a que se tendrá un porcentaje de no respuesta (real) que no conocemos hasta el día de la jornada electoral, una vez concluida la elección y con toda la información disponible se determinará la precisión realmente obtenida.
2. Los resultados de las simulaciones presentadas son únicamente para el método de estimación usado, y tienen como objetivo fijar un tamaño de

muestra. Sin embargo, se pueden emplear métodos de estimación alternativos que arrojarán precisiones distintas.

3. Se debe señalar que hasta el momento (abril de 2021) se cuenta con la proyección de 402 CAE (información proporcionada por la DEOE), de los cuales se ha estimado que en promedio 215 apoyarían el CR para la elección a la Gubernatura en el Estado de Tlaxcala. De esta estimación de CAE, el 73.42% (158) solo tendrá que reportar 1 sola casilla. Por lo cual, es pertinente señalar que el tamaño de muestra de 270 casillas se encuentra dentro de los límites de la capacidad operativa de campo.
4. Hay que mencionar, por un lado, que si se recibe un tamaño de muestra menor a 270 casillas, el día de la Jornada Electoral, todavía se pueden tener estimaciones confiables si la distribución de la muestra recibida es adecuada. Sin embargo, las precisiones serán menores como se puede ver en la Tabla 9. Por ejemplo, aún para un tamaño de muestra de 180 casillas con una composición y distribución suficiente desde el punto de vista teórico y que corresponda al Diseño muestral aquí definido, se alcanzaría una precisión muy cercana al 1.05%, pero se podría tener un sesgo importante en las estimaciones reportadas. Por otro lado, también se debe señalar que si se recibe una muestra menor a 180 casillas con una composición y calidad muy alejada de la aquí definida por el Diseño muestral, difícilmente se podrán obtener resultados que ofrezcan certidumbre en las estimaciones.
5. Finalmente hay que señalar que los diferentes procedimientos de estimación que implementarán los tres miembros del COTECORA del grupo 1 (Mtra. Claudia Ortiz, Dr. Alberto Alonso y Dr. Gabriel Núñez), para la realización del CR del Estado de Tlaxcala, estarán basados en el Diseño muestral definido en esta Sección.

III. Integración de las estimaciones

Cada uno de los 3 miembros del grupo 1 del COTECORA, con la misma información de las casillas recibidas el día de la jornada electoral, realizarán una estimación para la votación de cada uno de los candidatos a la Gubernatura del Estado de Tlaxcala, así como, para el porcentaje de participación. Por lo anterior, para integrar los resultados anteriores se propone usar la unión de al menos dos de las tres estimaciones obtenidas. Sin embargo, la manera definitiva

de integrar las estimaciones dependerá de los resultados que se obtengan en ejercicios de simulación y en los simulacros.

5.2.1.2. Equipo 2

5.2.1.2.1. Chihuahua, Colima, Nayarit y Zacatecas³⁷

Diseño muestral

El objetivo del conteo rápido es estimar la proporción de votos en favor de cada candidato. Para cumplir esto, se selecciona una muestra aleatoria de casillas del total y, con la muestra, se calculan intervalos del 95% de confianza para la proporción de votos de cada candidato. A continuación, explicamos el criterio bajo el cual se coleccionará dicha muestra aleatoria para Chihuahua, Nayarit, Zacatecas y Colima.

El diseño muestral será estratificado en donde, al interior de cada estrato, se seleccionarán casillas mediante un muestreo aleatorio simple sin reemplazo. La asignación de muestra será proporcional al tamaño de cada estrato. Para definir los estratos a emplear:

1. Se crean distintas posibles estratificaciones a partir de cruces de algunas variables.
2. Para cada estado, para cada una de las estratificaciones y para distintos tamaños de muestra, se simulan 200 muestras utilizando como marco la tabla de cómputos distritales del 2018.
3. Con estas muestras se estiman los intervalos de 95% de confianza para la proporción de votos.
4. Se evalúan ciertos criterios con el fin de elegir la estratificación más conveniente.

Los puntos importantes por definir entonces son: ¿Qué estratificaciones se evaluarán? ¿Qué tamaños de muestra se evaluarán? ¿Cómo se estimarán las proporciones de votos? ¿Con qué criterios decidiremos la estratificación óptima?

³⁷ Elaborado por la Dra. Michelle Anzarut Chacalo y la Mtra. María Teresa Ortiz Mancera con la colaboración de Dr. Luis Felipe González Pérez y la Mtra. Ixchel Meza Chávez.

Estratificaciones evaluadas

Utilizamos cuatro variables para construir estratificaciones:

- Distrito federal dentro de cada estado.
- Índice de marginación a nivel sección (1 o 2 subgrupos): Esta variable se crea realizando un análisis de componentes principales con los datos del censo de población y vivienda del 2010 [1] y utilizando la primera componente, que se interpreta como un índice de marginación.
- Tamaño de lista nominal (1 o 2 subgrupos).
- Índice de votos en 2015 a nivel sección. (1 o 2 subgrupos): Esta variable se crea realizando un análisis de componentes principales con los datos de votación del 2015 y utilizando la primera componente.

En cada estado, cruzamos el número de distritos federales con una de las siguientes configuraciones:

Número de estratificación	Índice de marginación	Lista nominal	Índice de votos
1	1	1	1
2	1	1	2
3	1	2	1
4	1	2	2
5	2	1	1
6	2	1	2
7	2	2	1
8	2	2	2

Colapsamos algunos estratos del diseño para conseguir un mínimo de 60 casillas por estrato, usando las variables de la tabla anterior de derecha a izquierda y sin colapsar distritos federales. El corte de las variables en subgrupos se hace de forma que cada subgrupo tenga el número más similar de casillas posible.

Tamaños de muestra evaluados

Mediante el muestreo probabilístico es posible definir estrategias de selección, tamaños de muestra y estimadores, para asegurar que la diferencia entre el estimador \hat{p} y la verdadera proporción p no difiera en valor absoluto en más de una cierta cantidad d con un 95% de confianza. A la cantidad d se le conoce como precisión. Las precisiones indican entonces el tamaño de los intervalos de confianza que presentaremos.

El tamaño de muestra a elegir está directamente relacionado con el nivel de confianza que queremos obtener, las precisiones deseadas y el tamaño de la población total. Buscamos el tamaño mínimo de muestra que nos permita obtener intervalos del 95% de confianza con buenas precisiones. En cuanto a las poblaciones totales, el estado de Chihuahua, en las elecciones del 2018, tenía 5286 casillas, Colima tenía 950, Nayarit 1668 y Zacatecas 2508. Considerando esto, para el estado de Chihuahua evaluaremos muestras del 4%, 6%, 8% y 10%, mientras que para los estados de Colima, Nayarit y Zacatecas, que tienen un número menor de casillas, evaluaremos muestras del 10%, 15%, 20%, 25% y 30%. Como las fracciones de muestreo consideradas para Chihuahua son distintas presentaremos su análisis por separado.

Método de estimación

Sea p_j la proporción de votos a favor del candidato j , para estimar p_j utilizaremos el estimador de razón combinado \hat{p}_j , definido como:

$$\hat{p}_j = \frac{\sum_{h=1}^L N_h \left(\frac{1}{n_h} \sum_{i=1}^{n_h} x_{hij} \right)}{\sum_{j=1}^J \sum_{h=1}^L N_h \left(\frac{1}{n_h} \sum_{i=1}^{n_h} x_{hij} \right)},$$

donde, para cada estrato h , n_h es el tamaño de la muestra, N_h es el número total de casillas, L es el número total de estratos, J es el número total de candidatos (incluyendo nulos y no registrados) y x_{hij} es el total de votos a favor del candidato j en la casilla i .

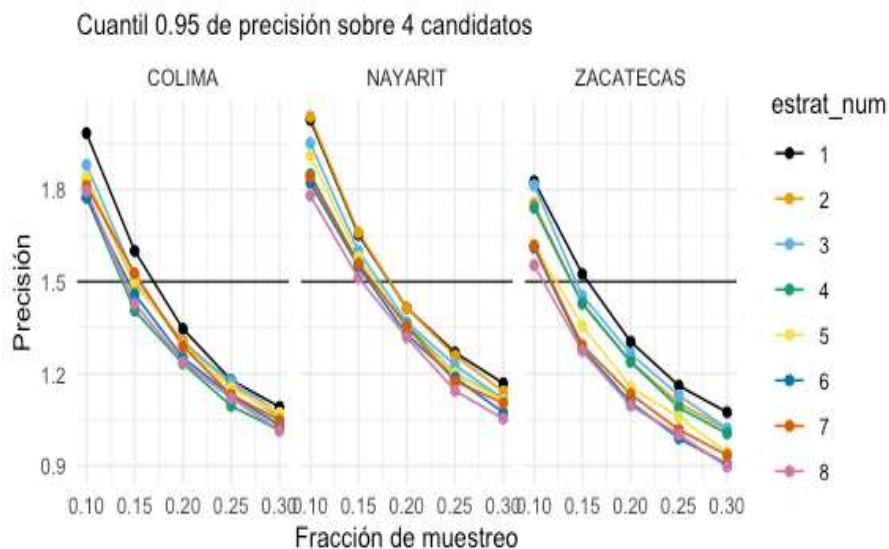
El error estándar de este estimador se calcula utilizando bootsap no paramétrico [2].

Criterios de decisión para la estratificación

Entre más fina es una estratificación, mejor es la precisión. Sucede lo mismo con el tamaño de muestra, entre mayor es la muestra mejor es la precisión. Sin embargo, utilizar estratificaciones demasiado finas causa que haya estratos con pocas casillas, que pueden no ser bien representados a la hora de muestrear. Por otro lado, el considerar muestras demasiado grandes implica un mayor porcentaje de capacitadores-asistentes electorales (CAE) con más de una casilla a reportar, esto complica el trabajo de los CAE y puede resultar en que no reporten alguna de las casillas. Por esta razón, el diseño muestral evalúa en conjunto estos dos criterios, con el objetivo de encontrar un balance que maximice la precisión sin poner en riesgo la obtención de la muestra.

Resultados de simulaciones para definir el tamaño de muestra para Colima, Nayarit y Zacatecas

A continuación, tenemos una gráfica que muestra el cuantil 0.95 de las precisiones sobre los 4 candidatos mayores para cada estado, para distintos tamaños de muestreo y para cada estratificación evaluada. La línea negra indica que el cuantil 0.95 de las precisiones es igual a 1.5, esto quiere decir que, bajo esa línea, el 95% de las muestras arrojaron intervalos con precisiones menores a 1.5. Un criterio conservador de selección es buscar fracciones de muestreo y estratificaciones que estén por debajo de la línea negra.



Notamos que para estos 3 estados la estratificación 6 tiene buena precisión, muy similar a la mejor precisión obtenida con la estratificación 8 que es la más fina, por lo tanto, se elige la estratificación 6 para los 3 estados.

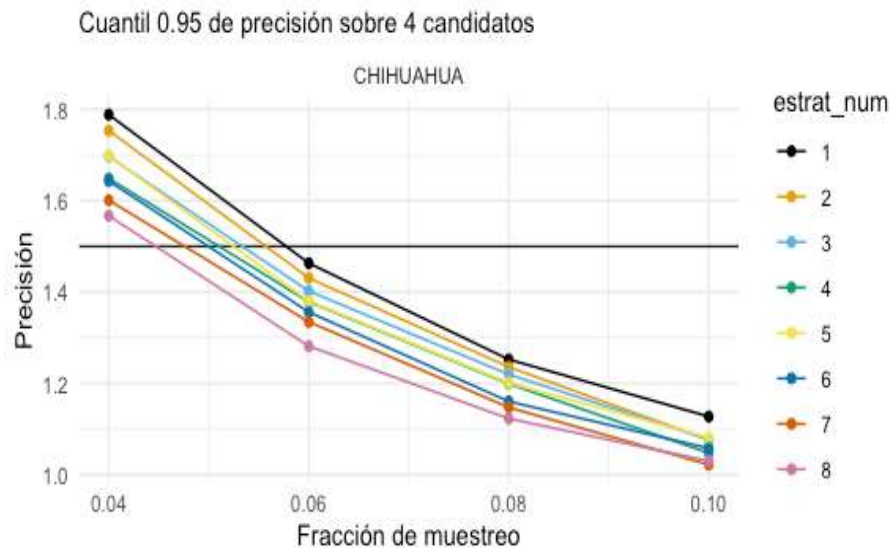
En cuanto a fracciones de muestreo tenemos que para Colima y Nayarit la fracción de 20% nos asegura estar bajo la línea negra, mientras que para Zacatecas basta tomar 15% con la estratificación 6. En la siguiente tabla mostramos el promedio de porcentaje de CAE con 1, 2 y 3 casillas para la estratificación 6 y algunas fracciones de muestreo que parece razonable analizar:

Estado	Fracción de muestreo	1	2	3
COLIMA	0.15	76	21	3
COLIMA	0.20	68	26	5
COLIMA	0.25	61	31	7
NAYARIT	0.15	73	23	4
NAYARIT	0.20	64	29	6
NAYARIT	0.25	56	32	10
ZACATECAS	0.15	78	20	2
ZACATECAS	0.20	71	25	4
ZACATECAS	0.25	63	30	6

El número de CAE con 1, 2 y 3 casillas parece razonable para fracciones de muestreo de 20% para Colima y Nayarit y 15% para Zacatecas, por lo que confirmamos ésta elección.

Resultados de simulaciones para definir el tamaño de muestra para Chihuahua

En Chihuahua observamos un patrón similar en cuánto a precisión:



En este caso, como las fracciones de muestreo consideradas son mucho menores, el porcentaje de CAE con 1, 2 y 3 casillas es razonable para cualquiera de estas fracciones.

Estado	Fracción de muestreo	1	2	3
CHIHUAHUA	0.04	94	6	0
CHIHUAHUA	0.06	90	9	0
CHIHUAHUA	0.08	87	12	1
CHIHUAHUA	0.10	84	15	1

Elegimos utilizar la estratificación número 3 con una fracción de muestreo de 6%.

Diseño final

Para el estado de Chihuahua se elige la estratificación 3 que cruza los distritos federales con 2 niveles de lista nominal. Para los estados de Colima, Nayarit y Zacatecas se elige la estratificación 6 que cruza los distritos federales con 2 niveles de índice de marginación y 2 niveles de índice de votos. A continuación, aparece el diseño muestral final.

Estado	Fracción de muestreo	Número de estratos	Tamaño de muestra
CHIHUAHUA	0.06	18	317
COLIMA	0.20	6	190
NAYARIT	0.20	10	334
ZACATECAS	0.15	14	376

Referencias

[1] <http://gaia.inegi.org.mx/geoelectoral>

[2] https://tereom.github.io/est_computacional/05-Bootsrap.html

5.2.1.2.2. Michoacán³⁸

Objetivos.

El objetivo del Conteo Rápido Michoacán 2021, para la elección de Gobernador, es obtener estimaciones del porcentaje de votos emitidos a favor de los candidatos a Gobernador del estado, así como del porcentaje de ciudadanos que acudan a votar. El Conteo Rápido es un ejercicio estadístico y para llevar a cabo se desarrollarán diversos procedimientos técnicos que tienen sustento en la Teoría de la Inferencia Estadística.

Antecedentes.

La Inferencia Estadística se ha desarrollado progresivamente durante más de 100 años y a partir de su estructura conceptual, el rigor de sus resultados teóricos y su desempeño en la práctica se ha establecido como una disciplina científica muy reconocida y como un estándar aceptado internacionalmente para el estudio de los fenómenos que presentan variabilidad. En el caso específico de este conteo rápido se recurrirá, por una parte, a la Teoría de Muestreo para Poblaciones Finitas, con el propósito de establecer, en primer lugar, un diseño eficiente de la muestra que proveerá la información para la estimación. Por otra parte, de entre los Métodos Bayesianos de Estimación Paramétrica, se seleccionará el mecanismo que habrá de utilizar con los datos que ofrezca la muestra para producir las estimaciones.

Diseño Muestral.

Consideraciones generales.

Para efectos de este conteo rápido se considera como población bajo estudio a la totalidad de las casillas que se instalarán en la entidad el día de la jornada electoral. En esa población es de interés estimar, por una parte, el número total de votos que se recibirán para la elección de Gobernador y consecuentemente, el porcentaje de ciudadanos, en la lista nominal, que acudirán a ejercer ese

³⁸ Elaborado por el Dr. Manuel Mendoza Ramírez con la colaboración del Mtro. Gerardo Orantes Jordan

derecho. Además, y como objetivo central, interesa también estimar el porcentaje de votos emitidos a favor de cada candidato a Gobernador.

Estas estimaciones se producirán a partir de la información recolectada, al cierre de la votación, en una muestra de casillas que se habrá de definir mediante un mecanismo de selección probabilística la cual, con la finalidad de atender a lo dispuesto en el Artículo 373 del Reglamento de Elecciones del Instituto Nacional Electoral, debe abarcar la mayor dispersión geográfica posible. Para cumplir con ese propósito, se utilizará un diseño muestral estratificado. De esta forma, la entidad se dividirá en regiones que servirán como la base para definir los estratos y de cada uno de los estratos se seleccionará una muestra de casillas completamente al azar, sin reemplazo. El tamaño de la muestra en cada estrato será proporcional al tamaño del propio estrato.

Este diseño, además de responder al requerimiento del reglamento, es muy común en la práctica del muestreo. Produce estimaciones más precisas que las correspondientes al caso en que la muestra se selecciona completamente al azar y sin reemplazo del conjunto de todas las casillas instaladas en la entidad sin restricción, geográfica o de otro tipo. Para definir un diseño estratificado es necesario establecer, en primer lugar, los criterios con que se dividirá la entidad en estratos y, una vez con tales criterios, el siguiente paso es seleccionar el tamaño de total de la muestra.

La determinación de los criterios de estratificación se lleva a efecto tomando en cuenta que el diseño produce mejores resultados en la medida en que las unidades dentro de cada estrato son más homogéneas entre sí y, al mismo tiempo, diferentes de las que pertenecen a otros estratos. La homogeneidad en esta conceptualización, se refiere al comportamiento de la variable, objeto de estudio; en este caso, la votación observada.

Por otra parte, el tamaño de la muestra se fija de manera que la estimación alcance niveles aceptables de confiabilidad y precisión. En el caso del Conteo Rápido para la elección de Gobernador en Michoacán en 2021, el nivel de confiabilidad, como también lo establece el Reglamento de Elecciones, se ha fijado en 0.95. De esta forma, para cada propuesta de estratificación, es necesario examinar la precisión que producen los distintos tamaños de muestra, considerando este nivel de confiabilidad.

Criterios de estratificación.

Para un conteo rápido estatal una posibilidad, relativamente natural, es considerar como posible estratificación la división de la entidad en sus distritos electorales. En México, existen dos tipos de distritos electorales: los distritos locales y los distritos federales. Los primeros son la base para los procesos electorales que sólo incumben a la entidad como elecciones de Gobernador y del Congreso Estatal. Los segundos, juegan el mismo papel pero para los procesos federales, en particular para la elección de Presidente de la República y de la Cámara de Diputados Federal.

En primera instancia podría considerarse la división de la entidad en sus distritos electorales locales. Sin embargo, El proceso electoral del año 2021 incluye una multiplicidad de elecciones simultáneas y, en particular, el conteo rápido en Michoacán forma parte de un sistema de 16 conteos rápidos simultáneos que desarrollará el Instituto Nacional Electoral.

Este sistema incluye otros 14 conteos estatales para las respectivas elecciones de gobernador y un conteo rápido nacional para estimar la integración de la Cámara de Diputados Federal. Estos ejercicios de estimación no solamente se han de llevar a cabo de manera simultánea, sino que deben aprovechar todas las sinergias posibles, para garantizar el más alto nivel de eficiencia operativa.

En estos términos, y en virtud de que el conteo rápido para la Cámara de Diputados necesariamente utilizará los 300 distritos electorales federales como base para la estimación, es conveniente definir inicialmente la estratificación en Michoacán, para la elección de Gobernador, a partir de los 12 distritos electorales federales en que se divide el estado.

Adicionalmente, se evaluaron otros criterios que permiten segmentar a la entidad con división aún más fina. En este punto es conveniente comentar que, si bien una división con un mayor número de estratos potencialmente puede producir estimaciones más precisas, este no es siempre el caso. El resultado depende, como ya se mencionó, de que se cumplan las condiciones de homogeneidad dentro de estratos y heterogeneidad entre estratos. Por otra parte, un mayor número de estratos implica que el número de casillas por estrato disminuya y esta condición puede conducir a que no se cuente con información de todo un estrato el día de la jornada electoral si la muestra llega incompleta, como ya ha ocurrido en procesos electorales previos. Así, es necesario también establecer una restricción sobre el tamaño mínimo de los estratos resultantes en el diseño.

En cualquier caso, para tomar la decisión sobre la estratificación en el caso de Michoacán, como criterio adicional al de los distritos electorales federales, se consideraron los distritos electorales locales tomando ventaja de que, a diferencia de lo que ocurre en otros estados de la república, en esta entidad cuando se considera la distritación federal más reciente, ocurre que los distritos locales constituyen un refinamiento de los distritos federales. En concreto, cada distrito federal se divide exactamente en dos distritos locales.

Ahora bien, también se exploró el criterio que divide los distritos en sus componentes urbana y no urbana. Además, se analizó la división de cada distrito en dos segmentos, según el tamaño del listado nominal de las casillas que lo integran (un segmento formado por las casillas con un tamaño de listado nominal mayor al tamaño mediano en el distrito y otro segmento con el resto de las casillas).

Finalmente, se examinó también una división de cada distrito en dos grupos de casillas según el valor de un índice de marginación específicamente definido en la zona en que se encuentra cada casilla (un grupo incluye a las casillas con índice de marginación mayor al valor mediano en el distrito y el otro grupo incluye el resto de las casillas).

De esta manera en el caso de Michoacán, se consideraron distintas propuestas de estratificación a partir de estos criterios: Distrito Electoral Federal (DF), Distrito Electoral Local (DL), Tipo de Sección (URB), Tamaño del Listado Nominal (LN) y Valor del Índice de Marginación (IM). Como ya se indicó el criterio de partida fue el distrito electoral federal, para considerar divisiones de esos distritos respecto a los otros criterios. Si bien se exploraron diversas combinaciones con uno, dos o tres criterios simultáneamente, al final, con el propósito de que el tamaño de los estratos no se redujera en exceso, sólo se consideraron cinco alternativas.

Se trata de la estratificación con los Distritos Federales solamente (DF) y de las cuatro que se obtienen al cruzar los distritos federales con cada uno de los otros cuatro criterios: Distritos Locales (DFxDL), Tipo de Sección (DFxURB), Listado Nominal (DFxLN) e Índice de Marginación (DFxIM).

Evaluación de las alternativas de estratificación.

Una vez establecidas las cinco estratificaciones posibles, el procedimiento para seleccionar la alternativa óptima hizo uso de los resultados de la elección local para Gobernador de 2015. A partir de la base de datos que contiene los

resultados de la votación en cada una de las casillas instaladas en esa jornada electoral, y con el conocimiento de los resultados finales, se procedió, para cada estratificación propuesta y considerando distintos tamaños de muestra, a evaluar la precisión de las estimaciones sujeta a un nivel de confiabilidad de 0.95.

De esta manera, para una estratificación fija y un tamaño total de muestra, que se distribuye entre los estratos en forma proporcional al tamaño de cada estrato, se procede de la siguiente manera.

1. Se selecciona una muestra con ese tamaño de acuerdo con el esquema de estratificación.
2. A partir de la información de esa muestra, se calcula la distribución final para la proporción de votos a favor de cada uno de los candidatos a gobernador.
3. Para el i -ésimo candidato, se calcula la estimación puntual de la proporción de votos a su favor \hat{p}_i y, con el valor finalmente registrado en esa elección, p_i , se registra el correspondiente error de estimación, así como el error de estimación $|p_i - \hat{p}_i|$.
4. También para el i -ésimo candidato, se calcula el intervalo de estimación de \hat{p}_i con un nivel de confiabilidad del 95% y se registra la longitud de ese intervalo.
5. Considerando los resultados de todos los candidatos con esa muestra se reporta M , el máximo de los errores de estimación y L , el valor máximo de las longitudes de los intervalos de estimación.

Los pasos 1 a 5 se repiten un número K de veces, a fin de estudiar el comportamiento de los indicadores M y L , además de monitorear la cobertura de los intervalos de estimación.

Con los resultados de este estudio es posible comparar, considerando un tamaño de muestra fijo, el desempeño de las distintas estratificaciones para elegir la alternativa óptima. Una vez con esta decisión, se puede estudiar el efecto de distintos tamaños de muestra para seleccionar uno que conduzca a un nivel de precisión que genere la certidumbre requerida.

En este estudio, como ocurrirá el día de la jornada electoral, la distribución a posteriori o final de cada proporción de votos y, por tanto, las estimaciones puntuales y por intervalo, se producen con el método que se consigna en el

apartado de Métodos de Estimación y se describe en detalle en Mendoza y Nieto (2016).

Resultados de la elección para Gobernador en 2015

En la elección de 2015 contendieron seis candidatos por el cargo de Gobernador del estado de Michoacán. (LMCH) Luisa María Calderón Hinojosa por el partido Acción Nacional, (MAO) Manuel Antúnez Oviedo por el partido Movimiento Ciudadano, (MLNR) María de la Luz Núñez Ramos por el partido Morena, (GDB) Gerardo Dueñas Bedolla por el partido Humanista, (AOB) Ascensión Orihuela Bárcenas por la coalición formada por el partido Revolucionario Institucional y el partido Verde, y finalmente, (SAC) Silvano Aureoles Conejo por la coalición formada por el partido de la Revolución Democrática, el partido del Trabajo, el partido Nueva Alianza y el partido Encuentro Social.

Estructura del estudio de estratificación

Como ya se señaló, se consideraron los siguientes criterios para la establecer la estratificación.

Criterio base:

- Distrito Electoral Federal (DF).

Criterios complementarios:

- Distrito Electoral Local (DL)
- Tipo de Sección (URB)
- Tamaño del Listado Nominal (LN)
- Valor del Índice de Marginación (IM)

La información correspondiente a los primeros cuatro criterios se obtuvo directamente de la base de datos de la elección para Gobernador de 2015. El índice de marginación proviene del estudio de Anzarut y Ortiz (2021) donde se analizaron los factores que podrían influir en el arribo parcialmente tardío de la

muestra en los ejercicios de conteo rápido realizados por el INE y también corresponde con la información vigente en 2015.

Con estos criterios, y después de considerar diversas combinaciones, se decidió centrar el interés en estratificaciones que combinaran el criterio base con cada uno de los criterios complementarios. Especialmente, se privilegió la idea de asegurar que no se produjeran estratos de un tamaño reducido en exceso. No es conveniente que los estratos tengan un número de casillas muy pequeño ya que, en caso de que se reciba una muestra incompleta, se corre el riesgo de que se pierdan estratos completos. En este estudio, se fijó un umbral de 100 casillas como tamaño mínimo para los estratos.

Por otra parte, tomando en cuenta la experiencia de los conteos rápidos para las elecciones de Gobernador en 2018, el estudio considera tamaños de muestra que correspondieran a fracciones de muestreo entre 5 y 7 por ciento con intervalos de medio punto porcentual. Finalmente, para cada evaluación se seleccionaron $K = 1000$ muestras. Es decir, cada combinación de estratificación y tamaño de muestra se evalúa con el desempeño observado en un conjunto de 1000 muestras de la población de interés.

Así, se tienen cinco estratificaciones en el estudio:

1. Distrito Electoral Federal (DF)
2. Distrito Electoral Federal y Distrito Electoral Local (DFxDL)
3. Distrito Electoral Federal y Tipo de Sección (DFxURB)
4. Distrito Electoral Federal y Tamaño del Listado Nominal (DFxLN)
5. Distrito Electoral Federal e Índice de Marginación (DFxIM)

Finalmente, es importante consignar que con la estructura distrital de 2015, los distritos locales no resultan en un refinamiento de los correspondientes distritos federales. Esta característica, que simplifica la definición de los estratos resultantes de esta combinación de criterios, sólo se observa si se consideran los distritos federales de 2020 que serán los vigentes en la elección de 2021. Tomando en cuenta esta condición, se decidió considerar en el estudio las cinco estratificaciones con la distritación federal vigente en 2015 para replicar exactamente lo sucedido en la elección de ese año, pero también las cinco estratificaciones donde los distritos federales se definen de acuerdo con la más reciente distritación. De esta manera, en estricto sentido se evaluaron 10 estratificaciones.

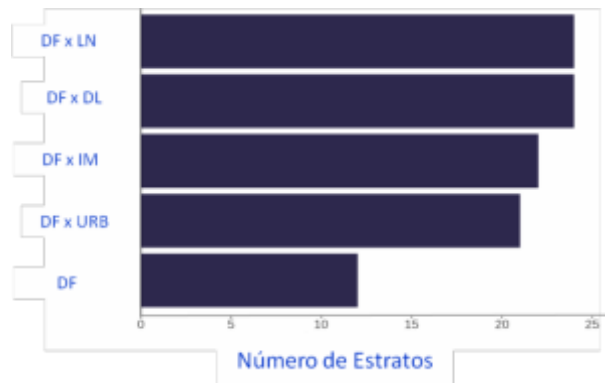
Resultados del estudio de estratificación

Como se ha indicado, el análisis de las posibles estratificaciones se realizó tanto con los distritos federales de 2015 como con los distritos de 2020. Una diferencia menor que se aprecia entre estas dos alternativas es en el número de estratos que se obtienen para las combinaciones de criterios. En la Figura 1 se muestra el número de estratos cuando se utiliza la distritación vigente en 2015 mientras que la Figura 2 exhibe lo propio con la información de 2020.

Figura 1. Número de estratos 2015



Figura 2. Número de estratos 2020



El segundo aspecto que se reporta es la cobertura que tienen los intervalos de estimación que nominalmente se producen con una confiabilidad de 95%. Estos resultados se reportan para cada una de las estratificaciones consideradas. En realidad, lo que verifica esta fase del estudio, con las muestras completas, es que

el modelo que se utiliza para producir las inferencias describe en forma apropiada la información muestral. Por esa razón, basta con llevar a cabo el análisis sólo para un tamaño de muestra. En este caso se utilizó un tamaño equivalente a una fracción de muestreo de 6%, el punto medio de las fracciones consideradas en el estudio. Los resultados se exhiben en las Figuras 3 y 4.

Figura 3. Cobertura de los intervalos de estimación 2015.

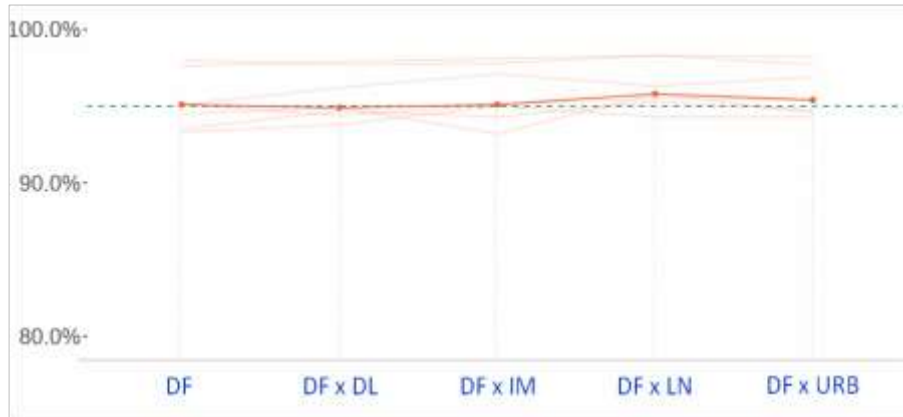


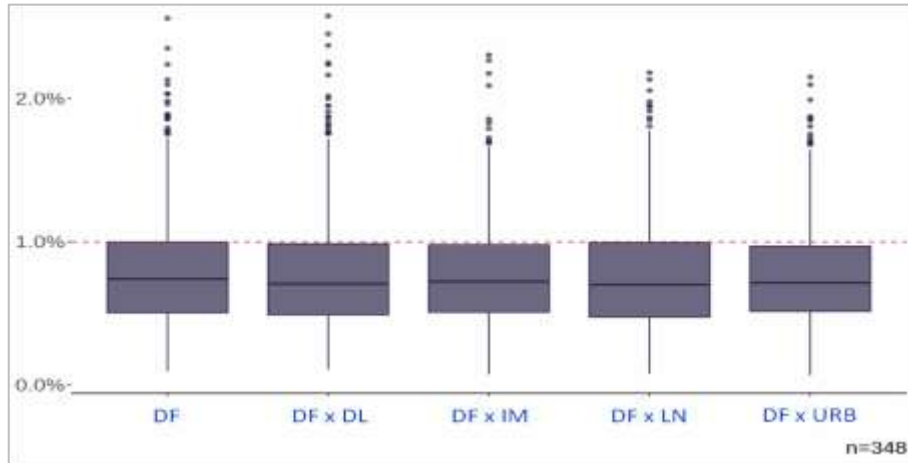
Figura 4. Cobertura de los intervalos de estimación 2020.



En las gráficas se observa que las coberturas para los intervalos de estimación correspondientes a todos los candidatos se encuentran alrededor del valor nominal 0.95 para todas las estratificaciones y se reporta también la cobertura mediana en el grupo de candidatos en cada caso. Los resultados coinciden con lo esperado.

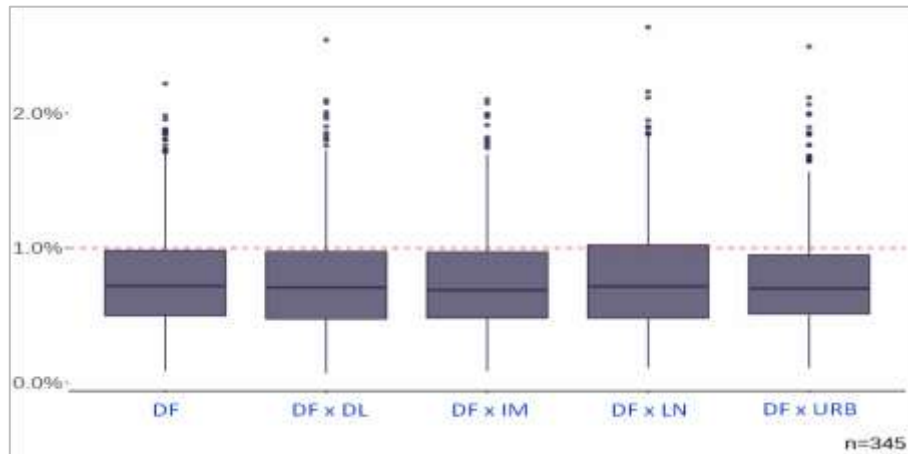
El tercer aspecto que se analizó fue el comportamiento del error máximo de estimación. Como se indicó previamente, para cada muestra se calculan los errores de estimación para los seis candidatos y lo que se analiza el máximo de esos errores a lo largo de las 1000 repeticiones. Este análisis también se produjo con un tamaño de muestra equivalente a una fracción de muestreo de 6%.

Figura 5. Error máximo de estimación 2015.



En las gráficas, se puede observar que sólo hay diferencias muy menores entre el comportamiento que presenta el error de estimación en las diferentes estratificaciones y este fenómeno se presenta tanto con la distribución federal de 2015 como con la de 2020.

Figura 6. Error máximo de estimación 2020.



También es notable que con esta fracción de muestreo en 75% de los casos el error **máximo** se mantiene por debajo de un punto porcentual. El siguiente elemento de análisis en el estudio, se refiere al máximo de las longitudes de los intervalos de estimación.

Figura 7. Longitud máxima del intervalo de estimación 2015.

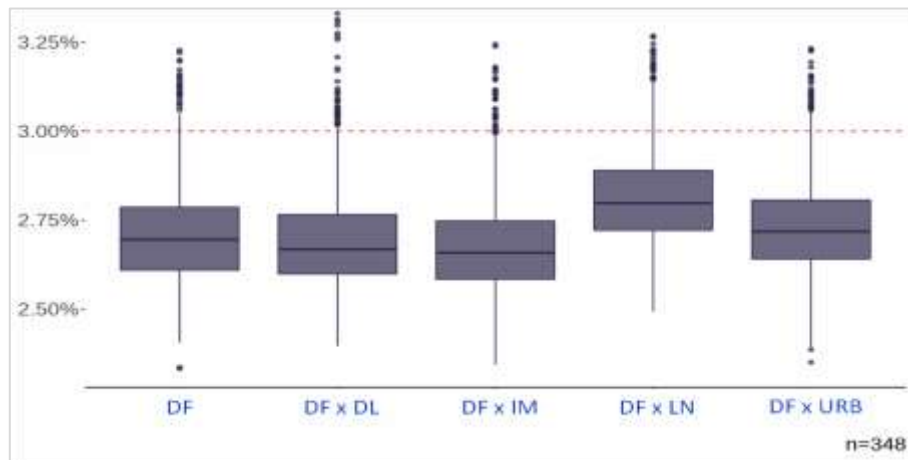
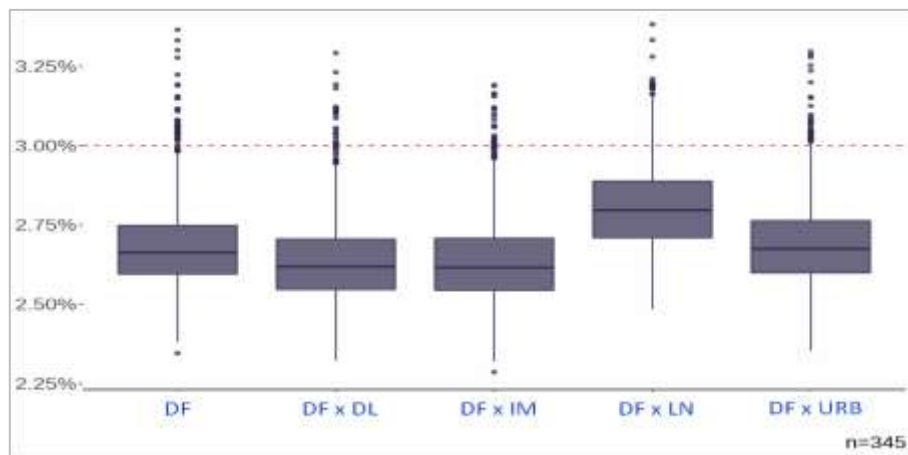


Figura 8. Longitud máxima del intervalo de estimación 2020.



En este caso nuevamente se consideran los intervalos para los seis candidatos en cada muestra y se calcula la longitud máxima. Aquí se mantuvo como referencia la fracción de muestreo de 6%.

De las gráficas se desprende que, nuevamente, el comportamiento de este indicador es similar con las dos distritaciones federales (2015 y 2020). En prácticamente todos los casos se observa una longitud máxima que no supera los tres puntos porcentuales. Este resultado se puede interpretar en el sentido de que con una fracción de muestreo de 6% se espera que, en el peor de los casos, los intervalos quedasen definidos como la estimación puntual más o menos 1.5 puntos porcentuales.

Ahora bien, la única estratificación que exhibe un patrón ligeramente distinto al resto es la que cruza distritos federales con el tamaño del listado nominal. Para esta estratificación se presenta un ligero incremento en la longitud máxima de los intervalos. Este es un tema que cobra importancia a la luz del análisis de Anzarut y Ortiz (2021) que estudia los factores que pueden influir en el arribo tardío o incluso en el no arribo de una parte de la muestra.

Estas autoras encontraron que un tamaño del listado nominal grande puede influir en el retraso de la información. De hecho, en este estudio se incorporó ese criterio lo mismo que el índice de información como consecuencia de los hallazgos en el trabajo de Anzarut y Ortiz. En su trabajo muestran que el patrón de arribo tardío o no arribo en las muestras de los conteos rápidos que realizó el INE en 2018 no suele ser aleatorio y que, en algunos casos, ese retraso induce un sesgo en las inferencias de manera que, por ejemplo, los intervalos de estimación pueden ser aparentemente precisos (de una longitud breve) pero no capturan el verdadero valor de la proporción de interés en virtud del sesgo.

En el estudio que aquí se trata, se incluyeron el tamaño del listado nominal y el valor del índice de marginación en primer lugar con la idea de que si existen estratos que específicamente incluyen a las casillas que son susceptibles de llegar en forma tardía o no llegar para el cómputo de las estimaciones, el monitoreo en arribo de la muestra debería de mostrar un desbalance en el caso de esos estratos como una señal de advertencia antes de intentar la producción de las inferencias con la muestra incompleta disponible.

En segundo término, hemos observado que, en particular, el tamaño del listado nominal, cuando se cruza con el distrito electoral federal, y se considera la muestra completa, no modifica el error máximo de estimación, pero aumenta ligeramente la longitud del intervalo con que se estima. Este puede ser un elemento particularmente útil que debe examinarse con muestras incompletas. De corroborarse este efecto con muestras incompletas que presenten un patrón de retraso similar al de conteos rápidos previos, este efecto podría mitigar, al menos parcialmente, el potencial sesgo debido a la llegada de una muestra incompleta.

Como conclusión de esta sección estudio se tiene que dos estratificaciones que podrían ser consideradas son la que únicamente utiliza como estratos los doce Distritos Electorales Federales (DF) y la que cruza los Distritos Electorales Federales con las dos categorías del Tamaño del Listado Nominal (DFxLN).

Determinación del tamaño de muestra.

Para fijar el tamaño de muestra apropiado, tomando en cuenta los resultados de la sección previa, se consideró la estratificación base (DF) con la distritación de 2020 que estará en vigor en 2021 y se explora el desempeño de las inferencias con muestras completas que corresponden a distintas fracciones de muestreo. Con 1000 muestras de cada tamaño se evalúan los indicadores M , el máximo de los errores de estimación y L , el valor máximo de las longitudes de los intervalos de estimación. Los resultados se pueden observar en a Figuras 9 y 10.

Figura 9. Error de estimación máximo como función del tamaño de la muestra.

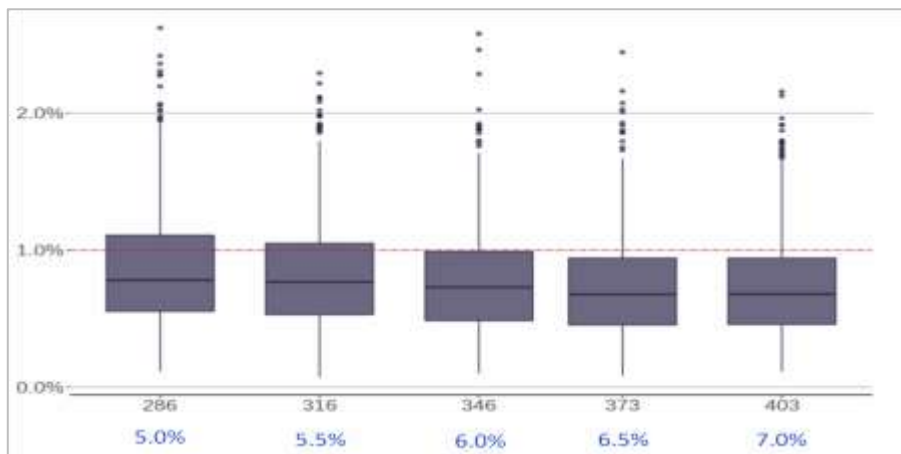
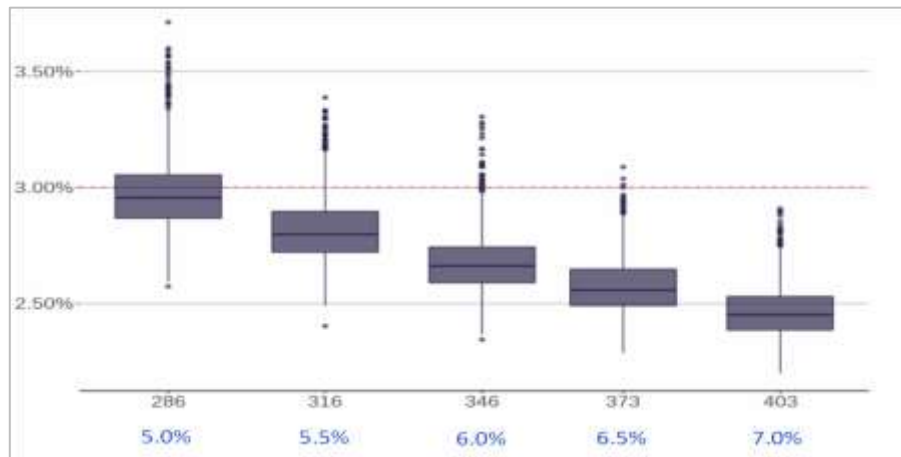


Figura 10. Longitud máxima del intervalo como función del tamaño de la muestra.



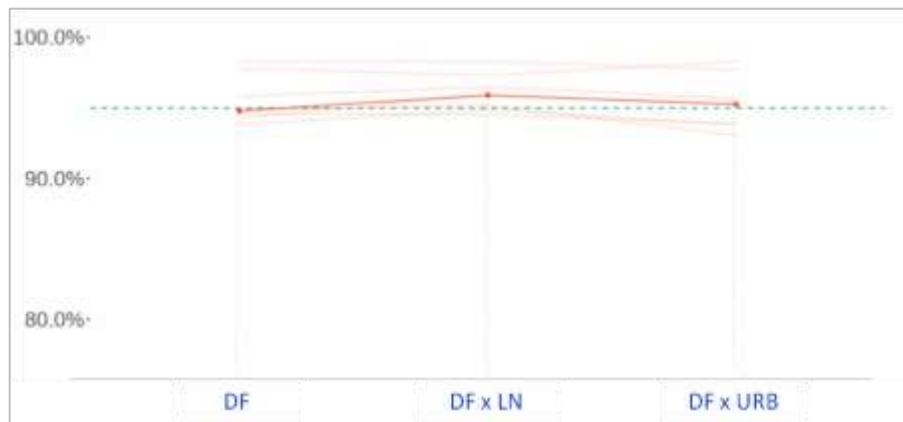
En estas gráficas se puede observar claramente que el error de estimación máximo disminuye a medida que aumenta el tamaño de la muestra, aunque no se producen cambios drásticos. En cualquier caso, a partir de una fracción de muestreo de 6.0% se tiene que al menos el 75% de las veces el error de muestreo **máximo** es inferior a un punto porcentual (con una muestra completa). Por otra parte, en la Figura 10, se aprecia como el tamaño de la muestra tiene un efecto más pronunciado en la longitud máxima de los intervalos de estimación. Aquí se observa que con una fracción de muestreo de 6.0% y 7.0%, la longitud **máxima** se encuentra por debajo de 3 puntos porcentuales. Esto significa que, en el peor de los casos, con una muestra completa, el resultado de la estimación se podría reportar como la estimación puntual más menos 1.5 puntos porcentuales. Como resultado final de esta sección se concluye que respecto al tamaño de muestra una fracción entre 6.0% y 7.0% es razonable.

Análisis con muestras incompletas.

En esta sección se analiza el desempeño del procedimiento de estimación partiendo de una muestra de tamaño equivalente a una fracción de muestreo de 6.0% que llega incompleta. Para establecer el patrón de censura se utilizó el modelo de Anzarut y Ortiz (2021) de manera que después de seleccionar una muestra de casillas completa de la población de interés, se les asignó una hora de arribo con el modelo indicado y el análisis se realizó únicamente con las casillas que hubieran llegado hasta las 22:30. Se produjeron 1000 repeticiones independientes del mecanismo y se reportan los resultados de Cobertura, M el máximo de los errores de estimación y L , el valor máximo de las longitudes de los

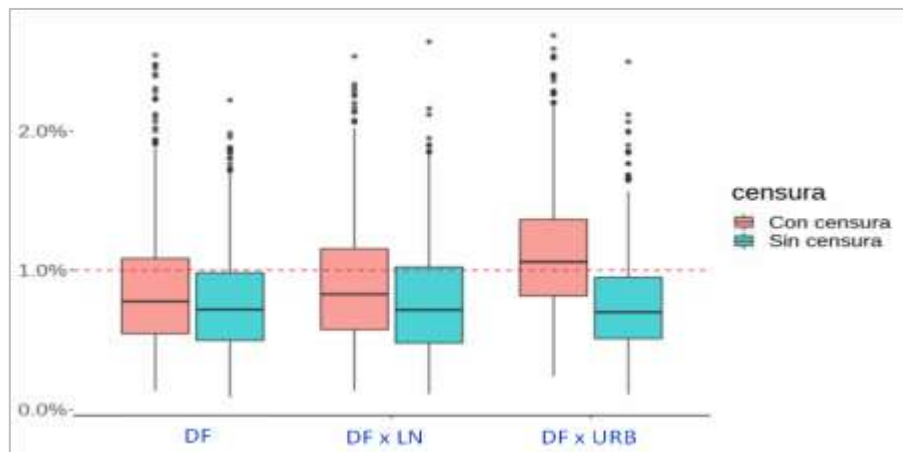
intervalos de estimación. Se reportan los resultados para las dos estratificaciones consideradas útiles (DF) y (DFxLN) y, como complemento, la estratificación (DFxURB). Los resultados se exhiben en las Figuras 11, 12 y 13.

Figura 11. Cobertura de los intervalos de estimación con muestras incompletas.



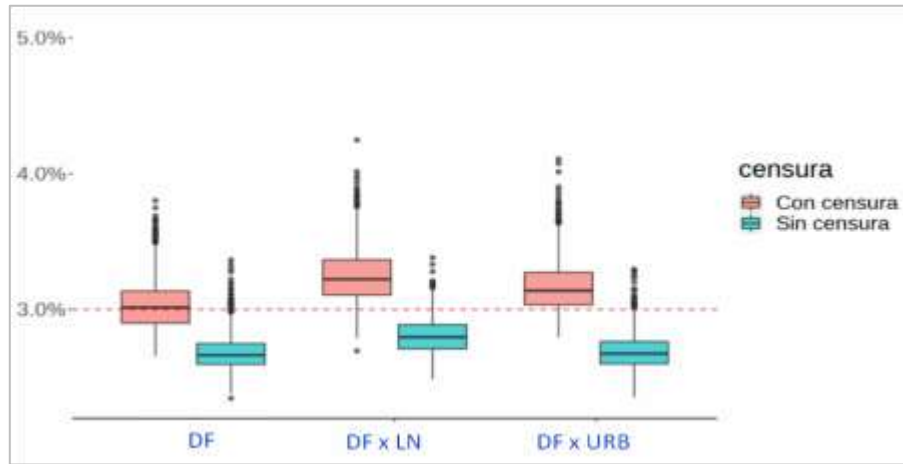
Como puede constatarse en la Figura 11, la cobertura nominal de 95% no se ve afectada por la censura de la muestra.

Figura 12. Error de Estimación Máximo con muestras incompletas.



Cuando se analiza el error máximo de estimación (Figura 12) se observa que en el caso de las dos estratificaciones candidatas este error aumenta, pero el aumento no es de gran magnitud. En cambio, con la estratificación (DFxURB), el error de estimación máximo se incrementa considerablemente más.

Figura 13. Longitud Máxima del intervalo de estimación con muestras incompletas.



Finalmente, respecto a la longitud máxima del intervalo de estimación, se observa que la estratificación (DFxLN) es la que reacciona más frente a la censura de la muestra ampliando los intervalos. Esta precisamente es la que se puede considerar una medida defensiva frente al problema de eventual sesgo atribuible a la censura. De esta forma, el análisis sugiere utilizar la estratificación que cruza los Distritos Electorales Federales con las dos categorías del tamaño de Listado Nominal.

Carga de trabajo de los Capacitadores Asistentes Electorales.

Un elemento que también es objeto de análisis en el proceso para determinar el tamaño de la muestra para los conteos rápidos es la carga de trabajo que la recogida de la muestra representa para los Capacitadores Asistentes Electorales (CAE). Específicamente, interesa que por su tamaño la muestra no implique que un porcentaje muy alto de los CAE a cargo de recoger y transmitir la información de la muestra tengan que reportar los resultados de dos o más casillas.

En esta sección se utiliza la estratificación base (DF) 2020. Se seleccionan muestras de la población de interés para cada fracción de muestreo, desde 4% hasta 7%, con intervalos de medio punto porcentual. En cada muestra, se analizan las áreas de responsabilidad electoral y se calcula el porcentaje de CAE con sólo una casilla en muestra. También se reporta también el porcentaje de CAE con dos y más de dos casillas en muestra.

El resultado se puede observar en la Figura 14 y 15. En la primera se puede observar que con una fracción de muestreo de 6.5%, prácticamente en un 100% de los casos, ocurre que al menos un 88% de los CAE que reportan información para el conteo rápido tienen a su cargo solamente una casilla.

Figura 14. Porcentaje de CAE con una casilla a su cargo como función del tamaño de muestra.

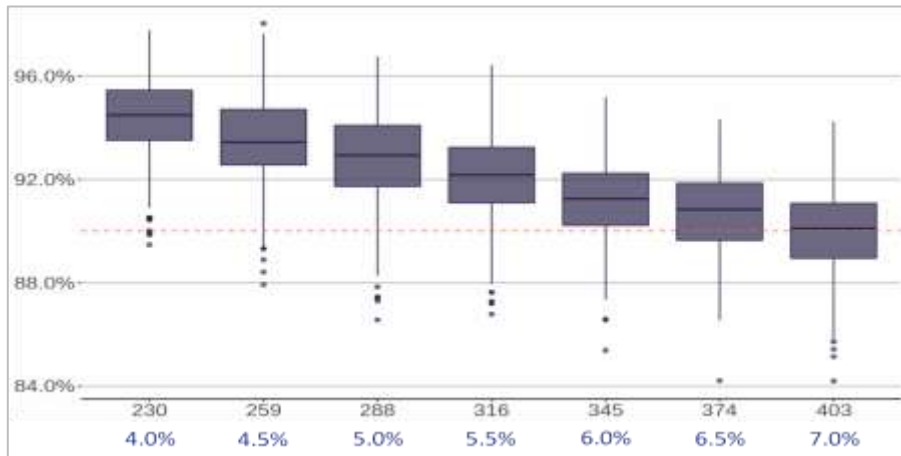
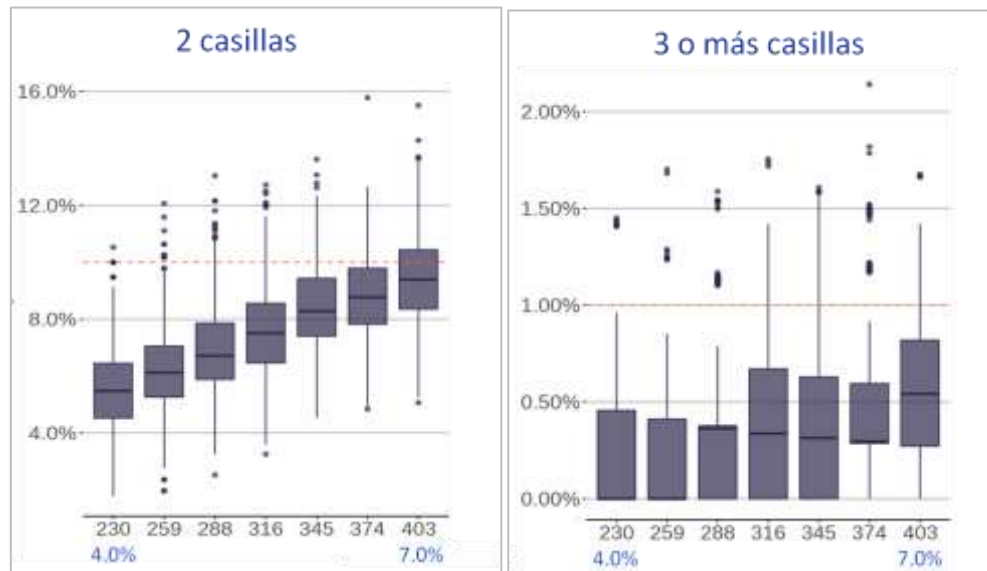


Figura 15. Porcentaje de CAE con más de una casilla a su cargo como función del tamaño de muestra.



Por su parte, de las gráficas en la Figura 15 se tiene que con una fracción de muestreo de 6.5% en más de un 75% de los casos, el porcentaje CAE con dos casillas se encuentra por debajo del 10% y el porcentaje de CAE con tres o más casillas en prácticamente el 100% de los casos se encuentra por debajo de 0.6%. Todos estos niveles son aceptables.

Consideraciones Finales.

A partir de los resultados obtenidos se propone utilizar la estratificación que cruza los Distritos Electorales Federales con dos categorías del tamaño del Listado Nominal. El resultado es un conjunto de 24 estratos.

La variable Listado Nominal tiene un efecto menor en el desempeño de las estimaciones si se cuenta con las muestras completas. Sin embargo, con muestras que llegan con patrón de atraso como el de 2018, este factor incrementa la longitud de los intervalos de estimación y esta característica puede operar como una medida para mitigar el posible sesgo debido a la censura.

Respecto al tamaño de la muestra, se considera suficiente un factor de muestreo de 6.5% que para la elección de 2020 se espera se traduzca en, aproximadamente, 400 casillas.

Referencias.

Anzarut, M. y Ortiz, T. (2021). Análisis del patrón de arribo tardío de la muestra de los conteos rápidos del Instituto Nacional Electoral. Reporte Interno del Comité Técnico para los Conteos Rápidos 2021.

Mendoza, M. y Nieto-Barajas, L.E. (2016). Quick counts in the Mexican presidential elections: A Bayesian approach. *Electoral Studies* 43, 124-132.

5.2.1.3. Equipo 3

5.2.1.3.1. Baja California, Campeche, Querétaro, San Luis Potosí y Sonora³⁹

Diseño para la selección de la muestra.

En todos los estados se considera un diseño estratificado aleatorio con asignación proporcional de la muestra a los estratos con el objetivo de tener una buena cobertura geográfica

Determinación del tamaño de muestra y estratificación

Para definir la conformación de los estratos a utilizar, se considera la información de las elecciones de gobernador del 2015 en todos los estados excepto en Baja California donde se utiliza la información de las elecciones de gobernador de 2019.

Se estudian varias estratificaciones:

1. Sin estratos
2. Distritos federales como estratos
3. Cruce de Distritos Federales con tipo de sección
4. Cruce de Distritos Federales con municipios

Al evaluar estas estratificaciones se cuida en primera estancia, que los estratos formados tengan un número razonable de casillas de donde obtener la muestra.

Para determinar los tamaños de muestra, para cada tipo de estratificación, se estima la proporción de votos para cada candidato, utilizando el estimador de razón combinado y su aproximación al error cuadrático medio por series de Taylor, correspondiente a un muestreo estratificado aleatorio con asignación proporcional de la muestra a los estratos.

³⁹ Elaborado por la Mtra. Adriana Ducoing Watty, la Mtra. Patricia Romero y el Dr. Raúl Rueda con la colaboración del Dr. Guadalupe Jiménez, el Lic. Enrique Reyes y el Mtro. Javier Santibáñez

Se elige la estratificación más eficiente en términos de tamaño de muestra y precisión elegida (0.8% o 0.9%) para los dos candidatos que presentan la mayor variabilidad, respetando que se tenga una confianza del 95% y se elige el tamaño de muestra mayor.

La determinación del tamaño de muestra para cada candidato se realiza del siguiente modo:

$$n = \frac{z_{(1-\frac{\alpha}{2})}^2 * N \sum_{h=1}^L N_h S_h^2}{X^2 \delta^2 + z_{(1-\frac{\alpha}{2})}^2 * \sum_{h=1}^L N_h S_h^2}$$

Donde:

δ : precisión deseada

X: total de votos emitidos en la entidad

L: Número de estratos

N: Total de casillas en la entidad

N_h : Total de casillas en el estrato h de la entidad

$z_{(1-\alpha/2)}$: cuantil de la distribución normal estándar para una confianza de (1- α)

$$S_h^2 = \frac{\sum_{i=1}^{N_h} [(Y_{hi}^p - \bar{Y}_h^p) - R^p (X_{hi} - \bar{X}_h)]^2}{N_h - 1}$$

Y_{hi}^p : total de votos en la casilla i del estrato h para el candidato p

\bar{Y}_h^p : promedio de votos en el estrato h para el candidato p

X_{hi} : total de votos emitidos en la casilla i del estrato h

\bar{X}_h : promedio de votos emitidos por casilla en el estrato h

R^p : proporción de votos a favor del candidato p en la entidad

Resultados obtenidos

Estado	Estratos	Tamaños de muestra
Baja California	8 Distritos Federales	667
Campeche	7 grupos de municipios dentro de DF	365
Querétaro	8 Tipo de sección en DF	385
San Luis Potosí	7 Distritos Federales	559
Sonora	7 Distritos Federales	410

Determinación del número de Casillas por ARE

Con el objetivo de explorar cuántas casillas de la muestra les tocarían a cada CAE con los tamaños de muestra obtenidos para cada entidad, se realizan simulaciones con 1000 repeticiones y se calcula el porcentaje de CAE a los que les toca 1, 2, 3 o 4 casillas.

Estado/ casillas	Porcentaje de CAE con una, dos, tres y cuatro casillas			
	1	2	3	4
Baja California	76.16	20.73	2.89	0.21
Campeche	53.65	33.25	10.84	2.04
Querétaro	75.26	21.12	3.3	0.32
San Luis Potosí	75.43	21.24	3.13	0.29
Sonora	78.69	18.75	2.38	0.17

Referencias

Lohr, Sharon L. (2010). "Muestreo: Diseño y Análisis". Thomson Learning. NY, USA.

Kish, L. (1970). "Survey Sampling". John Wiley and Sons. NY, USA.

Levy, Paul S. & Lemeshow, Stanley. (2008). "Sampling of Populations: Methods and Applications. 4th ed. Wiley. USA.

5.2.2. Procedimiento de estimación

5.2.2.1. Dr. Alberto Alonso y Coria⁴⁰

El procedimiento que a continuación se describe se utilizará para realizar las estimaciones de Baja California Sur y Tlaxcala

Procedimiento de estimación

Para estimar el porcentaje de votación que corresponde a cada partido con respecto a la votación total emitida y el porcentaje de participación ciudadana en la elección, tanto para el caso de Baja California Sur como de Tlaxcala, se utilizarán dos procedimientos clásicos de estimación para un muestreo aleatorio estratificado.

1. A través del porcentaje de votación con respecto al listado nominal.

Como primer paso se estima el número de votos que obtendría cada una de las fuerzas políticas y el total de votos emitidos en el estrato h.

Sea \bar{v}_i el promedio de votación para el candidato i en las casillas recibidas de la muestra correspondiente al estrato h y sea \bar{ln} listado nominal promedio de estas casillas. Si LN es el listado nominal real del estrato, el estimado para el total de votos para para el estrato h del candidato i está dado por:

$$\widehat{Y}_i = \widehat{R}_i * LN \text{ en donde } \widehat{R}_i = \frac{\bar{v}_i}{\bar{ln}}$$

El cálculo de la varianza para el estimador \widehat{Y}_i se obtiene de la siguiente forma:

$$V(\widehat{Y}_i) = \frac{N^2}{n} * (1 - f) * \left[\frac{\sum_{j=1}^n (v_{ij} - \widehat{R}_i * ln_j)^2}{n - 1} \right]$$

Donde $f = \frac{n}{N}$ es la fracción de muestra, N es el número total de casillas en el estrato h y n el número de casillas recibidas para ese estrato.

Para obtener la estimación de la votación obtenida por cada candidato a nivel estatal (v_{Ti}) y en razón de que los estratos son ajenos y cubren a todo el territorio,

⁴⁰ Con la colaboración del Fis. Joaquín Morales Bolío

simplemente se suma la estimación realizada para cada estrato. Un procedimiento semejante se utiliza para el cálculo de la varianza total.

Los intervalos de confianza para el total de votos recibidos correspondientes a un partido o alianza i estarán dados entonces por:

$$v_{Ti} \pm z_{\alpha/2} * \sqrt{V(\hat{Y}_{Ti})}$$

donde $1 - \alpha$ es el nivel de confianza establecido. Para la estimación se ha definido el nivel de confianza como 95% por lo que $z_{\alpha/2} = 1.96$ en razón de la correspondiente área bajo la curva de la ecuación normal.

De aquí se procede a realizar los cocientes correspondientes para obtener el porcentaje de votación con respecto a la votación total emitida.

2. A través de la razón de la votación con respecto a la votación emitida.

En este caso se construye el estimado puntual de votación recibida a través de:

$$\hat{Y}_{st,i} = \sum_h N_h * \bar{v}_{i,h}$$

El estimador combinado de razón estaría dado por $\hat{R}_{c,i} = \frac{\hat{Y}_{st,i}}{\hat{Y}_{st,total\ votos}}$

La varianza es ahora:

$$V(\hat{Y}_{c,i}) = \sum_h \frac{N_h^2}{n_h} * (1 - f_h) * [S_{vi,h}^2 + \hat{R}_{c,i}^2 * S_{vt,h}^2 - 2 * \rho_h S_{vi,h} S_{vt,h}]$$

En forma semejante al inciso anterior de aquí se procede a generar los intervalos de confianza para cada uno de los candidatos.

Referencias

Cochran, W. G. (1977) Sampling Techniques. 3ed. Wiley&Sons, Inc.

5.2.2.2. Dra. Michelle Anzarut Chacalo y Mtra. María Teresa Ortíz Mancera⁴¹

Procedimiento de estimación

El objetivo del conteo rápido es producir estimaciones de la proporción efectiva de votos a favor de cada uno de los candidatos a gobernador usando los votos registrados en una muestra. El modelo que utilizaremos para estimar dicha proporción es un modelo jerárquico bayesiano, dada la estructura jerárquica natural de los datos de observación, tiene sentido plantear un modelo en donde algunos coeficientes varíen por estrato, esto permite estimar coeficientes para grupos subrepresentados.

El modelo está basado en el modelo de postestratificación de [1], al considerar a las casillas postratificadas por algunas covariables, proporciona un medio para controlar los sesgos asociados a tales covariables, corrigiendo diferencias conocidas entre la muestra y las poblaciones objetivo. Esta característica es particularmente importante, pues se ha observado que en elecciones anteriores no se ha obtenido la muestra completa planificada a la hora de dar los resultados del conteo. Aún más, los datos que faltan por llegar no faltan completamente al azar, como consecuencia, la probabilidad de que no se informe una casilla puede depender de la respuesta que queremos medir.

Modelo

Veamos primero el modelo para la participación ciudadana. Sea y_i el número total de votos emitidos en la casilla i , suponemos que

$$y_i \sim \text{BinNeg}(\mu_i, \phi_i),$$

con media $\mu_i = n_i w_i$ y sobredispersión $\phi_i = \frac{n_i w_i}{\kappa_{s(i)}^{part}}$. El término n_i es el tamaño de la lista nominal en la casilla i , por lo que w_i representa la proporción de personas en la lista nominal que votaron. La varianza de y_i está dada por $n_i w_i \left(1 + \kappa_{s(i)}^{part}\right)$.

Ajustamos un modelo de regresión multinivel para el parámetro w_i .

⁴¹ Con la colaboración del Dr. Luis Felipe González Pérez y la Mtra. Ixchel Guadalupe Meza Chávez.

$$w_i = \text{invlogit}(\beta_{s(i)}^{\text{part},0} + \beta_{\text{part}}^T X_i)$$

donde X_i es una matriz de covariables centradas y estandarizadas para la casilla i .

Análogamente, modelamos el número de votos obtenidos para el candidato k en la casilla i como

$$y_{i,k} \sim \text{BinNeg}(\mu_{i,k}, \phi_{i,k}).$$

En este caso la media $\mu_i = n_i w_i \theta_{i,k}$ y la sobredispersión está dada por $\phi_i = \frac{n_i w_i \theta_{i,k}}{\kappa_{s(i)}}$, por lo que $\theta_{i,k}$ representa la proporción de personas que votaron por el candidato k del total de votos en la casilla i .

Sea p el número total de candidatos, el vector $\theta_i = (\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p})$ está dado por

$$\theta_i = \text{softmax}(\beta_{s(i)}^{0,1} + \beta_1^T X_i, \dots, \beta_{s(i)}^{0,p} + \beta_p^T X_i),$$

donde recordamos que

$$\text{softmax}(z_1, \dots, z_p)_j = \frac{\exp(z_j)}{\sum_{i=1}^p \exp(z_i)}$$

Tanto en el caso de la participación, como la proporción de votos para cada candidato, consideramos los coeficientes $\beta_{\text{part}}, \beta_1, \dots, \beta_p$ fijos sobre todo el estado, sin embargo, en ambos casos la ordenada al origen $\beta_{s(i)}^{\text{part},0}, \beta_{s(i)}^{0,1}, \dots, \beta_{s(i)}^{0,p}$ y el parámetro de sobredispersión $\kappa_{s(i)}^{\text{part}}, \kappa_{s(i)}$ tienen una estructura jerárquica dependiente del estrato, las distribuciones correspondientes pueden consultarse a detalle en [4]. Al añadir predictores a nivel estrato se reduce la variación no explicada dentro de cada estrato y, como consecuencia, se reduce la variación total, produciendo estimaciones más precisas.

Las covariables que utilizaremos son:

1. Tipo de sección (rural, urbana o mixta).
2. Índice de marginación.

Estimación de parámetros

Implementamos el modelo usando el software STAN. El método de estimación es completamente reproducible y puede encontrarse en el paquete de R 'quickcountmx' [3]. Con esto, estimamos la distribución de probabilidad del número total de votos por cada candidato en cualquier casilla del estado.

Construcción de estimaciones

La idea general de la estimación es la siguiente. Primero, para cada casilla que no está en la muestra, simulamos votos según el modelo. Después, agregamos los valores de las casillas observadas con los valores simulados, obteniendo votos para el total de las casillas. Con los votos es directo calcular la proporción de votos a favor de cada candidato y la participación ciudadana. Finalmente, los intervalos de probabilidad que presentaremos como resultado son una evaluación simple de estas distribuciones, es decir, cada intervalo es una región de 95% de probabilidad.

Para asegurarnos de poder cubrir una variedad de posibles resultados, hicimos una calibración bayesiana [4]. Esto implica que, aunque el modelo es bayesiano, se elige para producir inferencias bien calibradas en el sentido frecuentista. Específicamente, generamos 300 muestras de la base de datos de elecciones de 2015 y 2018, con distintos cortes en diferentes porcentajes de muestras observados y probamos si el modelo proporciona intervalos de probabilidad posteriores con aproximadamente su cobertura nominal.

La calibración nos permite elegir el conjunto óptimo de covariables, pero también demuestra que cuando tenemos una proporción considerable de casillas faltantes el modelo no basta para corregir los sesgos de las muestras censuradas. Como mencionamos, en los conteos rápidos que se han llevado a cabo una proporción considerable de las casillas (de 10% a 50%) está censurada por el tiempo. Esto implica que los parámetros estimados tendrán sesgo con respecto a los poblacionales.

Para corregir el sesgo haremos lo siguiente. Primero, estimamos la distribución de probabilidad del número total de votos por cada candidato en cualquier casilla del estado, después, a la hora de simular tomamos en lugar del vector θ_i una perturbación dada por

$$\theta'_i = \text{softmax}(\theta_{i,1} + \eta_1, \dots, \theta_{i,p} + \eta_p),$$

donde η_1, \dots, η_p son variables aleatorias independientes e idénticamente distribuidas $N(0, 0.1(1 - p_{obs}))$ y p_{obs} es el porcentaje de la muestra observado. La distribución inicial de η_1, \dots, η_p se fija de acuerdo con el sesgo observado en las elecciones de 2018, notemos que se simula un parámetro η_k para cada candidato k , entre mayor es η_k más crece el tamaño de los intervalos producidos y conforme nos aproximamos a la muestra completa η_k se aproxima a 0.

El modelo, con el cambio anteriormente propuesto, nos provee de una forma consistente de ajustar el sesgo y tratar los datos faltantes. Como resultado, sus coberturas son cercanas al 95% tanto con muestras completas como con muestras censuradas. Para tener un método adicional de respaldo utilizamos también el estimador de razón combinado, donde la varianza se calcula por remuestreo no paramétrico, este es un método tradicional de estimación de muestreo.

Referencias

- [1] Park, D.K., Gelman, A., Bafumi, J.: Bayesian multilevel estimation with poststratification: state-level estimates from national polls. *Political Analysis* 12(4), 375–385 (2004)
- [2] Anzarut, M., González, L. F., & Ortiz, M. T. A Multilevel Mixture Model accounting for sample bias for the Quick Count Mexican Elections of 2021. (2021)
- [3] quickcountmx: an R package to estimate election results from Mexico. <https://github.com/cotecora-team-2/quickcountmx> (2021)
- [4] Little, R.J.: Calibrated Bayes, an alternative inferential paradigm for official statistics. *Journal of official statistics* 28(3), 309 (2012)

5.2.2.3. Mtra. Adriana Margarita Ducoing Watty⁴²

Se estima la proporción de votos a favor de cada candidato a la gubernatura mediante el estimador de razón combinado de un muestreo estratificado con selección de una muestra aleatoria simple sin reemplazo en cada estrato y con asignación proporcional de la muestra a los l estratos:

$$\hat{R}_p = \frac{\hat{Y}_p}{\hat{X}_p} = \frac{\sum_{h=1}^l \hat{Y}_{hp}}{\sum_{h=1}^l \hat{X}_h} = \frac{\sum_{h=1}^l N_h * \bar{y}_{hp}}{\sum_{h=1}^l N_h * \bar{x}_h}$$

Donde:

\hat{R}_p : estimador de la proporción de votos a favor del candidato p en la entidad

\hat{Y}_p : estimador del total de votos a favor del candidato p en la entidad

\hat{X}_p : estimador del total de votos emitidos en la entidad

\hat{Y}_{hp} : estimador del total de votos a favor del candidato p en el estrato h

\hat{X}_h : estimador del total de votos emitidos en el estrato h

\bar{y}_{hp} : estimador del promedio de votos por casilla a favor del candidato p,
en el estrato h

\bar{x}_h : estimador del promedio de votos emitidos por casilla en el estrato h

N_h : total de casillas en el estrato h

Mediante la técnica de remuestreo conocida como Bootstrap no paramétrico, que consiste en obtener entre 500 y 1000 muestras aleatorias simples con reemplazo de tamaño n_h de la muestra original de cada estrato, se construye la distribución por muestreo del estimador combinado \hat{R}_p . Para la construcción del intervalo de confianza al 95 % para la proporción de votos a favor de cada candidato, se escogen los cuantiles 2.5 % y 97.5% de dicha distribución.

⁴² Con la colaboración del Mtro. Guadalupe Jiménez Villanueva

Referencias:

Efron, B. (1979). "Bootstrap methods: Another look at the jackknife". The Annals of Statistics, Vol 7. No 1, pp. 1-26.

Efron, B (1981)." Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods". Biometrika, 68,3 pp.589-599

Davison, A.C. and D.V. Hinkley (1997). "Bootstrap Methods and their Applications. Cambridge, England: Cambridge University Press

Raj, Des. (1968). "Sampling Theory". McGRAW-Hill. NY, USA

Shao Jun. (2003). "Impact of the Bootstrap on Sample Surveys". Statistical Science, Vol 18, No. 2, pp. 191-198.

5.2.2.4. Dr. Manuel Mendoza Ramírez⁴³

Introducción.

La estadística Bayesiana es una alternativa a la estadística frecuentista para la producción de inferencias sobre los parámetros desconocidos de un modelo. Aborda los problemas de inferencia como problemas de decisión en ambiente de incertidumbre, y recurre a una teoría de la decisión con fundamento axiomático para proveer las soluciones. (Bernardo y Smith, 2000).

Tanto bajo el enfoque frecuentista como bajo el Bayesiano, el fenómeno de interés es sujeto de observación a través de mecanismos que garantizan que los datos que se recolectan forman muestras aleatorias. En consecuencia, el comportamiento, la variabilidad, de estas muestras se describe a través de modelos de probabilidad.

En este contexto, las características del fenómeno que interesa conocer se identifican con el valor de uno o más parámetros del modelo y es en esos términos que los procesos de inferencia se concentran en aumentar y precisar el conocimiento sobre los parámetros.

Desde la perspectiva Bayesiana, la incertidumbre sobre los parámetros también se describe a través de modelos de probabilidad. En este caso no se describe variabilidad alguna sino el estado de conocimiento que, en un momento determinado, tiene el científico sobre el fenómeno bajo estudio. Con diferentes modelos se puede reflejar desde un conocimiento relativamente preciso hasta un desconocimiento casi completo sobre el valor del parámetro. De esta forma, bajo este paradigma, tanto las variables aleatorias observables (la muestra) como los parámetros fijos (pero desconocidos) son descritos mediante modelos de probabilidad.

La cuantificación de la incertidumbre sobre un parámetro puede producirse antes de contar con una muestra. El conocimiento teórico, histórico o incluso empírico, previo a la recolección de la muestra da lugar a esta valoración que se refleja en un modelo o distribución inicial o a priori $f(\theta)$.

⁴³ Con la colaboración del Mtro. Gerardo Orantes Jordan

Si es posible obtener la información muestral $\underline{X} = (X_1, X_2, \dots, X_n)$, proveniente del modelo $f(x|\theta)$, el estado de conocimiento inicial es susceptible de actualizarse para combinar la información inicial con la que provee la muestra.

Esta actualización se lleva a cabo mediante el mecanismo que establece el teorema de Bayes. El resultado es el modelo final o posterior $f(\theta|x)$, que combina las dos fuentes de información a través de la relación

$$f(\theta|x) = f(x|\theta)f(\theta)/f(x).$$

A partir del modelo final se producen las inferencias de interés sobre el parámetro θ , de manera que es indispensable establecer las características de $f(x|\theta)$. Si bien la fórmula de Bayes es una expresión matemática simple, en la práctica, la caracterización analítica de la distribución final se puede complicar debido, por ejemplo, a las dificultades en el cálculo de la constante de normalización $f(x)$. Gracias a los avances computacionales recientes y en especial a los algoritmos de simulación Monte Carlo vía cadenas de Markov (MCMC), es posible conocer las características de cualquier distribución final mediante métodos de simulación sin necesidad de calcular la constante de normalización (Chen et al. 2000).

Definiciones específicas

El objetivo del conteo rápido es producir estimaciones confiables de la proporción efectiva de votos a favor de cada uno de los candidatos. Esta proporción se calcula como el cociente entre el número de votos a favor de un candidato y el número total de votos emitidos por todos los candidatos incluyendo los votos nulos y los candidatos no registrados.

Respecto a la información muestral con la que se producen los resultados del conteo rápido, se adopta el supuesto de que corresponde con un diseño estratificado de casillas. En esos términos, sea N el número de estratos, cada uno con una lista nominal de tamaño n_i para $i = 1, \dots, N$. Además, se define θ_{ij} como la proporción de personas en la lista nominal del estrato i cuya preferencia es por el candidato j , donde $j = 1, \dots, J$ con J el número total de candidatos incluyendo los candidatos no registrados, los nulos y el abstencionismo. En este caso, la proporción de personas en la lista nominal de todo el estado cuya preferencia es por el candidato j está dada por

$$\theta_j = \sum_{i=1}^N \frac{n_i}{n} \theta_{ij}$$

donde n es el tamaño de la lista nominal del estado, si se trata de una elección local o la lista nominal del país si se trata de una elección federal. Finalmente, la proporción efectiva de votos a favor del candidato se puede calcular como

$$\lambda = \frac{\theta_j}{\sum_{i=1}^{J-1} \theta_i}$$

El Modelo y el Mecanismo de Inferencia

Tanto el modelo que se utiliza como el mecanismo para producir las inferencias se presentan en detalle en Mendoza y Nieto-Barajas (2016). En términos generales, se pueden describir como sigue. Del total de casillas en el estado (o país, según sea el caso) se seleccionan, mediante un mecanismo aleatorio c , de ellas (el tamaño total de la muestra), de manera que c_i de ellas provienen del i -ésimo estrato $i = 1, \dots, N$. De esta forma, c_i es el tamaño de la muestra en el estrato i .

Sea ahora, X_{ij}^k el número de personas a favor del candidato j , en la casilla k del estrato i . En esta formulación, $k = 1, \dots, K$ con K_i el número total de casillas en el estrato i .

El modelo que se emplea asume que

$$X_{ij}^k | \theta_{ij}, \tau_{ij} \sim N \left(n_i^k \theta_{ij}, \frac{\tau_{ij}}{n_i^k} \right)$$

donde n_i^k es el tamaño de la lista nominal de la casilla k en el estrato i , y τ_{ij} es un parámetro de precisión para $k = 1, \dots, c_i$, $i = 1, \dots, N$ y $j = 1, \dots, J$.

El modelo se complementa con la elección de las distribuciones iniciales para los parámetros θ_{ij} y τ_{ij} . En este caso, se utilizan iniciales que reflejan poca información a priori (distribuciones iniciales no informativas). El objetivo es obtener la distribución final para cada uno de los parámetros λ_j , con $j = 1, \dots, J - 1$ de manera que las inferencias que interesan se producen a partir de ese modelo de probabilidad. Como en otros casos, en este modelo esa distribución final es analíticamente compleja así que se recurre a métodos de simulación para obtener los resultados.

Referencias

Bernardo, J.M. y Smith, A.F.M. (2000). *Bayesian Theory*. Wiley: Chichester.

Chen, M.-H., Shao, Q.-M. e Ibrahim, J.G. (2000). *Monte Carlo methods in Bayesian computation*. Springer, New York.

Mendoza, M. y Nieto-Barajas, L.E. (2016). Quick counts in the Mexican presidential elections: A Bayesian approach. *Electoral Studies* **43**, 124-132.

5.2.2.5. Dr. Gabriel Núñez Antonio⁴⁴

Procedimiento de estimación

El reconocimiento explícito de la incertidumbre es crucial en las ciencias estadísticas. En muchas situaciones simples, la incertidumbre de un estimador puede visualizarse a través de cálculos analíticos basados en algún modelo probabilístico, adecuado. Sin embargo, en problemas más complejos esta tarea puede ser complicada, y los resultados llegan a ser potencialmente incorrectos si los supuestos básicos no se cumplen o si se realizan aproximaciones con la finalidad de simplificar los cálculos. Los avances computacionales permiten hoy en día el empleo de métodos denominados de remuestreo, los cuales ayudan a soslayar los problemas mencionados anteriormente, a la vez que permiten obtener estimaciones de errores estándar, intervalos de confianza, y otras medidas de incertidumbre.

La idea principal de los métodos *bootstrap* es el remuestreo a partir de los datos originales, ya sea en forma directa o vía un modelo ajustado; con la finalidad de obtener muestras replicadas a partir de las cuales se pueda evaluar la variabilidad de las cantidades de interés. Los métodos *bootstrap* también se pueden aplicar en problemas simples para verificar las características de las medidas de incertidumbre, para relajar supuestos, o para dar rápidas soluciones aproximadas. Un ejemplo de lo anterior es el remuestreo aleatorio para estimar la distribución permutacional de alguna estadística de prueba no paramétrica.

Es verdad que en muchas aplicaciones se puede confiar ampliamente en un modelo paramétrico particular y en el correspondiente análisis clásico basado en dicho modelo. Aun así, puede ser de utilidad investigar qué tanto se puede inferir sin asumir los supuestos de un modelo paramétrico particular. Esto es la esencia de la *robustez* del análisis estadístico realizado. El *bootstrap no-paramétrico* permite hacer esto. De acuerdo con lo anterior los métodos *bootstrap* se pueden aplicar tanto en los casos en los que se cuente con un modelo probabilístico bien definido para los datos, como en los casos en los que no se disponga de dicho modelo.

A pesar de su alcance y utilidad, el remuestreo debe ser aplicado en forma cuidadosa. A menos que las ideas básicas queden completamente claras, es muy fácil producir una solución al problema equivocado, o una mala solución al problema correcto. Los métodos de remuestreo como el *bootstrap* pretenden,

⁴⁴ Con la colaboración del Mtro. Emiliano Geneyro Esquarzón

de hecho lo hacen, evitar el empleo de procedimientos basados en supuestos distribucionales cuestionables.

La propuesta de estimación a emplear en los Conteos Rápidos del 6 de julio de 2021 para las Gubernaturas de los estados de Baja California Sur, Guerrero, Nuevo León, Sinaloa y Tlaxcala es utilizar estimadores de razón combinados usando métodos de remuestreo bootstrap no-paramétricos. Se debe señalar que en todos los casos se utilizarán los Diseños de muestreo definidos en cada una de las Secciones de este documento, correspondientes a los CR de cada estado. A continuación, se presentan las ideas básicas de los métodos de estimación propuestos.

Tipo de estimadores puntuales que se emplearán

Sea h el estrato a considerar, $h=1, \dots, k$, con N_h elementos en cada estrato. En cada estrato se toma una muestra aleatoria simple de tamaño n_h . Para estimar el porcentaje de votos de cada candidato p_j , $j=1, \dots, b$, los estimadores puntuales propuestos para la proporción de votos están dados por:

$$\hat{p}_j = \frac{\hat{T}_j}{\hat{T}} \quad \forall j = 1, \dots, b,$$

donde \hat{T}_j es un estimador estratificado del total de votos para cada candidato j , $j=1, \dots, b$, y \hat{T} es un estimador estratificado del total de votos obtenidos. Específicamente,

$$\hat{T}_j = \sum_h \frac{N_h}{n_h} Y_{j,h}, \quad \hat{T} = \sum_h \frac{N_h}{n_h} Y_h,$$

donde $Y_{j,h}$ es el total de votos del candidato j en el estrato h y Y_h es el total de votos en el estrato h .

Un procedimiento similar se utiliza para estimar la participación ciudadana en la elección correspondiente. Se debe mencionar que estos procedimientos de estimación ya los hemos empleado en los siguientes ejercicios de Conteo Rápido, los cuales han resultado exitosos:

1. El Conteo Rápido (CR) de 2015 para la elección federal de la conformación de la cámara de diputados.

2. El CR para la elección extraordinaria a la Gobernatura del Estado de Colima en 2016.
3. El CR para la elección a la Gobernatura del Estado de Nayarit en 2017.
4. Los Conteos Rápidos para el proceso electoral 2018.
 - a) Elección Presidencial.
 - b) Elección a la Jefatura de Gobierno de la CDMX.
 - c) Elección a la Gobernatura del Estado de Yucatán.
 - d) Elección a la Gobernatura del Estado de Veracruz.

(de la cual fui responsable directo)
5. Los Conteos Rápidos de 2019, para la elección a la Gobernatura de Baja California Sur y para la elección extraordinaria a la Gobernatura del Estado de Puebla.

Procedimiento para construir intervalos de confianza

Para obtener intervalos de confianza, con una precisión adecuada, para la proporción de votos para cada candidato se pueden seguir diferentes estrategias. Por ejemplo, se puede usar una aproximación normal asintótica. Si se considera que la aproximación normal asintótica puede no ser adecuada, otra opción es construir vía simulación la distribución muestral del estimador que se esté utilizando y obtener los cuantiles de orden 2.5% y 97.5%. Otra opción es, a partir de muestras bootstrap (no-paramétrico) obtener un estimador de la varianza del estimador correspondiente y con ellos construir el intervalo correspondiente. Estas dos últimas estrategias son las que se explorarían para construir los correspondientes intervalos de confianza, considerando procesos de remuestreo con al menos 5,000 simulaciones.

5.2.2.6. Mtra. Claudia Esther Ortiz Guerrero⁴⁵

Precisión en las estimaciones

El Conteo Rápido tiene el propósito de estimar las tendencias de la votación emitida por cada uno de los candidatos a la gubernatura del estado de Sinaloa. Para ello se ha determinado emplear el estimador de razón bajo un esquema de muestreo aleatorio estratificado

La estimación de la proporción poblacional es un promedio ponderado de los votos emitidos a cada candidatura sobre el total de votos de las muestras aleatorias en cada estrato, donde los pesos son los tamaños proporcionales al tamaño de los estratos.

No obstante, los resultados no se expresan como una estimación puntual, sino como intervalo de confianza, el cual consiste en dos valores que son los límites, inferior y superior, permitiendo acotar a un rango dentro del cual se encontrará la estimación puntual esperada con cierta probabilidad asociada.

Los intervalos de confianza, se estiman considerando, la precisión, la confiabilidad y la variabilidad en las elecciones. En medida de que exista menos variación, se podrán construir intervalos más estrechos que permiten dar una estimación más robusta, es decir, dar más certeza a que el resultado estimado refleje la tendencia real.

Para poder garantizar que las estimaciones se mantengan en los niveles de precisión deseados, la muestra total debe tener el suficiente número de elementos, elegidos aleatoriamente, tal que los resultados obtenidos reflejen correctamente los resultados de la votación.

Para determinar la precisión del estimador, se estimó su distribución de muestreo mediante un proceso de remuestreo tipo *Bootstrap* tomando con marco muestral los cómputos distritales de la elección a la gubernatura del estado de Sinaloa en 2015 y los cómputos distritales de la elección presidencial de 2018 asociados al estado de Sinaloa, basadas en 1,000 simulaciones.

La Técnica de *Bootstrap* es una de las que han resultado ser más eficientes para lograr un buen acercamiento del comportamiento en un número elevado de muestras. La idea principal del método *Bootstrap* es el remuestreo a partir de los

⁴⁵ Con la colaboración del Dr. Vicente González Juárez

datos originales, ya sea en forma directa o vía un modelo ajustado; con la finalidad de obtener muestras replicadas a partir de las cuales se pueda evaluar la variabilidad de las cantidades de interés.

5.2.2.7. Mtra. Patricia Romero Mares⁴⁶

Se estima la proporción de votos a favor de cada candidato a la gubernatura mediante el estimador de razón combinado de un muestreo estratificado con selección de una muestra aleatoria simple sin reemplazo en cada estrato y con asignación proporcional de la muestra a los l estratos:

$$\hat{R}_p = \frac{\hat{Y}_p}{\hat{X}_p} = \frac{\sum_{h=1}^l \hat{Y}_{hp}}{\sum_{h=1}^l \hat{X}_h} = \frac{\sum_{h=1}^l N_h * \bar{y}_{hp}}{\sum_{h=1}^l N_h * \bar{x}_h}$$

Donde:

\hat{R}_p : estimador de la proporción de votos a favor del candidato p en la entidad

\hat{Y}_p : estimador del total de votos a favor del candidato p en la entidad

\hat{X}_p : estimador del total de votos emitidos en la entidad

\hat{Y}_{hp} : estimador del total de votos a favor del candidato p en el estrato h

\hat{X}_h : estimador del total de votos emitidos en el estrato h

\bar{y}_{hp} : estimador del promedio de votos por casilla a favor del candidato p,
en el estrato h

\bar{x}_h : estimador del promedio de votos emitidos por casilla en el estrato h

N_h : total de casillas en el estrato h

El estimador de la varianza del estimador de la proporción de votos a favor del candidato p en la entidad es el siguiente:

$$\hat{V}(\hat{R}_p) = \frac{1}{\hat{X}^2} * \sum_{h=1}^l N_h^2 * \left(1 - \frac{n_h}{N_h}\right) * \frac{s_h^2}{n_h}$$

⁴⁶ Con la colaboración del Lic. Luis Enrique Reyes Romero

Donde:

$$s_h^2 = \frac{\sum_{i=1}^{n_h} [(y_{hi}^p - \bar{y}_h^p) - \hat{R}_p(x_{hi} - \bar{x}_h)]^2}{n_h - 1}$$

\hat{X} : total estimado de los votos emitidos en la entidad

y_{hi}^p : total de votos a favor del candidato p en la casilla i del estrato h

\bar{y}_h^p : promedio de votos a favor del candidato p por casilla en el estrato h

x_{hi} : total de votos emitidos en la casilla i del estrato h

\bar{x}_h : promedio de votos emitidos por casilla en el estrato h

n_h : número de casillas en muestra del estrato h

La precisión observada para la estimación del partido p se calcula como:

$$precision = z_{(1-\frac{\alpha}{2})} \sqrt{\hat{V}(\hat{R}_p)}$$

Considerando una confianza del 95% y el cuantil de la Normal estándar $z=1.96$, el intervalo para la proporción de votos del candidato p se obtiene de la siguiente manera:

$$(\hat{R} - precision, \hat{R} + precision)$$

Referencias:

Kish, L. (1970). "Survey Sampling". John Wiley and Sons. NY, USA.

Raj, Des. (1968). "Sampling Theory". McGRAW-Hill. NY, USA

Särndal, Carl Erik, et al. (1992). "Model Assisted Survey Sampling". Springer-Verlag. NY, USA.

5.2.2.8. Raúl Rueda Díaz del Campo⁴⁷

Se estima el total de votos emitidos a favor de cada partido

$$T_p = \sum_{i=1}^k \sum_{j=1}^{M_i} X_{ij},$$

en donde

$X_{ij} = (X_{ij1}, \dots, X_{ijp})$ son los votos emitidos en la casilla j del estrato i

M_i , el total de casillas en el estrato i

k , el número de estratos,

con base en una muestra de tamaño m , distribuida en los estratos $m = \sum_{i=1}^k m_i$.

Es fácil ver que

$$T_p = \sum_{i=1}^k (m_i \bar{x}_i + (M_i - m_i) \bar{X}_i^*),$$

con \bar{x}_i la media muestral observada en el estrato i y \bar{X}_i^* es la media **no** observada en el mismo estrato.

Bajo el supuesto de normalidad, no es difícil mostrar que la distribución predictiva de \bar{X}_i^* es una Student de dimensión p con parámetros

$$\left(m_i + 0.5, \bar{x}_i, \frac{(m_i - \nu + 1)(M_i - m_i)}{M_i} S_i^{-1} \right),$$

donde ν es el parámetro (conocido) de la distribución inicial y S_i la matriz de varianzas y covarianzas muestral.

Así que podemos obtener una muestra de esta predictiva, que es condicional a los datos observados, y con ella obtener una muestra de la distribución predictiva

⁴⁷ Con la Colaboración del Mtro. Javier Santibáñez Cortés

de T_p . Convertir estos totales a proporciones y estimar la participación es inmediato. El intervalo se construye con los cuantiles correspondientes.

Referencias

Bernardo, J.M. & Smith, A.F.M. (1994) Bayesian theory. Chichester: Wiley

O'Hagan, A. & Forstner, J. (2004). Bayesian inference. Kendall's Advanced Theory of Statistics Vol 2B. London: Arnold